To Move or to Convene:
Regulatory Circuits of Mat Fimbriae in *Escherichia coli*

Timo Lehti

Division of General Microbiology
Department of Biosciences
Faculty of Biological and Environmental Sciences
and
Viikki Doctoral Programme in Molecular Biosciences
University of Helsinki

ACADEMIC DISSERTATION

To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki in the auditorium 1041 at Viikki Biocenter 2 (Viikinkaari 5, Helsinki), on June 20th 2012, at 12 noon.

Helsinki 2012

Supervisors

Professor Timo K. Korhonen
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Docent Benita Westerlund-Wikström
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Advisory committee

Professor Jaana K. H. Bamford
Department of Biological and Environmental Science and Nanoscience Center
Faculty of Mathematics and Science
University of Jyväskylä, Finland

Professor Harri Savilahti
Department of Biology
Faculty of Mathematics and Natural Sciences
University of Turku, Finland

Reviewers

Professor Harri Savilahti
Department of Biology
Faculty of Mathematics and Natural Sciences
University of Turku, Finland

Docent Suvi Taira
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Opponent

Professor David Gally
Immunity and Infection Division
The Roslin Institute and Royal (Dick) School of Veterinary Studies
University of Edinburgh, United Kingdom

Custos

Professor Dennis Bamford
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

ISSN 1799-7372
ISBN 978-952-10-8026-5 (paperback)
ISBN 978-952-10-8027-2 (PDF); http://ethesis.helsinki.fi

Unigrafia
Helsinki 2012

To Nina and Milla

Contents

List of original publications

The thesis is based on the following articles (which are referred in the text by their Roman numerals) and unpublished results.

I **Timo A. Lehti**, Philippe Bauchart, Johanna Heikkinen, Jörg Hacker, Timo K. Korhonen, Ulrich Dobrindt, and Benita Westerlund-Wikström (2012). Mat fimbriae promote biofilm formation by meningitis-associated *Escherichia coli*. *Microbiology* **156**:2408-2417.

II **Timo A. Lehti**, Philippe Bauchart, Maini Kukkonen, Ulrich Dobrindt, Timo K. Korhonen, and Benita Westerlund-Wikström. Phylogenetic group –associated differences in regulation of the colonization factor Mat fimbria in *Escherichia coli*. Manuscript submitted for publication.

III **Timo A. Lehti**, Johanna Heikkinen, Timo K. Korhonen, and Benita Westerlund-Wikström (2012). The response regulator RcsB activates expression of Mat fimbriae in meningitic *Escherichia coli*. *Journal of Bacteriology* **194**, in press.

IV **Timo A. Lehti**, Philippe Bauchart, Ulrich Dobrindt, Timo K. Korhonen, and Benita Westerlund-Wikström (2012). The fimbriae activator MatA switches off motility in *Escherichia coli* by repression of the flagellar master operon *flhDC*. *Microbiology* **158**, in press.

Timo A. Lehti´s contribution to the articles:

I Laboratory experiments except for the initial observation of biofilm formation; participation in planning; drafting of the manuscript.

II Majority of the laboratory experiments, including strain construction, Northern blotting and mRNA half-life determination, primer extension, promoter-*lacZ* reporter analysis, protein-DNA interactions, immunological assays and sequence analysis; participation in planning; drafting of the manuscript.

III Laboratory experiments except for the initial screening of the transposon mutant library; participation in planning; drafting of the manuscript.

IV Laboratory experiments; participation in planning; drafting of the manuscript.

Summary

Escherichia coli is an extensively used model organism in molecular biology as well as an important human commensal and a versatile pathogen causing a repertoire of intestinal and extraintestinal infections. The pathogenic potential of *E. coli* is strongly influenced by laterally transferred genes, and pathogenic *E. coli* isolates are classified into distinct pathovars on the basis of the disease symptoms and the infection route. Molecular genetics and genomic studies have identified four main phylogenetic groups in the *E. coli* species; the extraintestinal isolates of *E. coli* cluster mainly into the B2 phylogroup, whereas diarrheagenic *E. coli* isolates are found in various phylogenetic groups. Distribution of virulence and fitness genes in *E. coli* phylogroups and pathovars has been extensively studied, but our knowledge on possible differences in gene regulation within specific groups of *E. coli* remains limited. It is however becoming apparent that evolutionary adaptation to pathogenic lifestyles involves, in addition to gain and loss of genes, changes in the expression of fitness genes found both in pathogenic and commensal *E. coli*. In this work, I investigated phylogenetic group -associated differences in promoter architecture and regulation of a conserved adhesin of *E. coli*, the Mat fimbria, and analyzed biological functions of this common organelle.

The Mat fimbria is encoded by the *mat* (or *ecp*) gene cluster, which is prevalent and conserved in the published genomes of *E. coli*. The biological functions of Mat and the regulatory mechanisms controlling the *mat* genes differ in *E. coli* isolates from different sources, and a basis for the regulatory differences was here described. This study identified the Mat fimbria as a novel determinant for temperature-dependent biofilm formation in *E. coli*. The *mat* operon was expressed *in vitro* in Luria-Bertani broth exclusively by the isolates of the B2 phylogroup, and the *mat* expression was induced by low growth temperature, by acidic pH, and by elevated levels of acetate. Extensive genetic studies and analysis of a well-described strain collection revealed that the differential *mat* expression is due to phylogroup-associated variation in the structure of the *mat* upstream regulatory DNA. Three regulatory genes, *matA*, *hns*, and *rcsB,* that participate in the transcriptional and post-transcriptional regulation of the *mat* operon were identified. Notably, these regulators form a network that reciprocally coordinates two opposite functions, Mat fimbriae-mediated adherence and flagella-driven motility.

The *mat* operon is an early acquisition in the evolution of *E. coli* and has been maintained during evolution of the *E. coli* groups. This thesis work presents a model for intra-species divergent evolution of regulatory DNA that supports phenotypic diversity within a bacterial species. The findings give an example of pathoadaptive variation in a regulatory DNA that controls expression of a common surface organelle of *E. coli*.

1 Introduction

1.1 *Escherichia coli* – a friend and a foe

The Gram-negative, rod-shaped bacterium *Escherichia coli*, originally called *Bacterium coli commune*, was first isolated from feces of a healthy infant in 1885 by Theodor Escherich (Escherich, 1885). *E. coli* has been the subject of intense research for over a century. Nowadays it is by far the best-studied bacterium and most extensively used model organism and laboratory workhorse in molecular biology. *E. coli* is an ecological generalist that inhabits various vertebrata including mammals, birds, reptiles, and fish (Clermont *et al.*, 2008; Escobar-Páramo *et al.*, 2006). However, *E. coli* is predominantly isolated from mammals (Gordon & Cowling, 2003) and in fact represents the most common member of the family *Enterobacteriaceae* in the mammalian lower intestinal tract (Gordon & FitzGibbon, 1999). *E. coli* is also widely distributed in the environment and capable of surviving in soil, water and sediment (Ishii *et al.*, 2006; Power *et al.*, 2005; Savageau M.A., 1983; Walk *et al.*, 2007). In humans, the majority of *E. coli* isolates are commensals of the lower intestinal tract (Tenaillon *et al.*, 2010), but *E. coli* also includes life-threatening, pathogenic strains that are responsible for diseases in the gut, urinary tract, bloodstream, and central nervous system (Croxen & Finlay, 2010; Kaper *et al.*, 2004).

The genus *Escherichia* belongs to the family *Enterobacteriaceae* of the *Gammaproteobacteria* that also includes *Salmonella*, from which *Escherichia* diverged about the time of the emergence of mammals, 120-160 million years ago (Ochman & Wilson, 1987). The genus consists of three named species, *E. coli* (the type species for the genus), *Escherichia fergusonii* (Farmer et al., 1985), and *Escherichia albertii* (Huys et al., 2003), as well as five phylogenetically distinct, cryptic clades that are genetically divergent from *E. coli* on the basis of multi-locus sequence typing (MLST) and whole-genome phylogenetic analysis (Luo *et al.*, 2011; Walk *et al.*, 2009). The available data indicate that the cryptic clades represent environmentally adapted bacteria that primarily live outside the host intestinal tract (Ingle *et al.*, 2011; Walk *et al.*, 2009). *Shigella* and *Escherichia* are generally treated as separate genera. The genetically very closely related *Shigella* spp. (*Shigella boydii*, *Shigella dysenteriae*, *Shigella flexneri*, and *Shigella soneii*) have adopted an intracellular lifestyle (Ray *et al.*, 2009) and are the causative agents of bacillary dysentery or shigellosis. Based on phylogenetic analyses, the invasive *Shigella* are specialized pathogenic strains of *E. coli* rather than separate species (Lan *et al.*, 2004; Pupo *et al.*, 1997; Pupo *et al.*, 2000). The genus status of *Shigella* has been retained largely because of their distinct clinical presentation and importance as pathogens.

E. coli strains are phenotypically heterogeneous, and isolates that share particular features have been divided into subgroups on various criteria. As early as in the 1940's, certain clinical manifestations were found to be associated with specific serological variants of *E. coli* (Bray, 1945). In 1947 Kauffman proposed an antigenic scheme for *E. coli* serotypes on the basis of expression of three types of surface antigens, the lipopolysaccharide (O), capsule (K), and flagellum (H) (Kauffmann, 1947). Since then, at least 173 O, 80 K, and 56 H antigens have been recognized (Ørskov & Ørskov, 1992), and several hundred *E. coli* serotypes have been described. However, only a small percentage of these serotypes has

been implicated in pathogenesis of *E. coli*. On the basis of distinct pathogenic schemes, pathogenic *E. coli* are separated from commesal isolates into different categories, pathovars, that are capable of causing disease in healthy individuals. Only to a limited extent, serotypic markers correlate with specific pathovars. In addition to serotyping, phylogenetic grouping of *E. coli* strains by multilocus enzyme electrophoresis, MLST or multiplex PCR are popular ways to categorize *E. coli* strains.

1.1.1 *E. coli* as a commensal

Human physiology and health is fully dependent on microbes that outnumber the total number of human cells in the adult body by tenfold (O'Hara & Shanahan, 2006). Dynamic interactions between the microbiota, the intestinal epithelium, and the host immune system maintain mucosal homeostasis and are essential for mounting protective immunity to invading microorganisms (Artis, 2008; Hooper & Macpherson, 2010). Microbes also provide metabolic capabilities important for the human host, such as cleavage of dietary polysaccharides, synthesis of short-chain fatty acids (SCFAs), and production of essential amino acids and vitamins. From the approximately 100 trillion (10^{14}) microbial cells in the human body, vast majority reside in the gastrointestinal tract (Savage, 1977). This microbial community, the human gut microbiota, is comprised of mostly bacteria, and recent metagenomic sequencing data from fecal samples of 124 European individuals suggest that this community is composed of more than 1,000 bacterial species (Qin *et al.*, 2010). The composition of human microbiota varies remarkably across body habitats (Costello *et al.*, 2009), including the different intestinal sites (Hayashi *et al.*, 2005), and from individual to individual (Eckburg *et al.*, 2005; Turnbaugh *et al.*, 2009).

Development of the microbiota commences during birth. From an essentially sterile state at birth, the body becomes rapidly colonized within hours by microbes that originate from the mother's vaginal, fecal and skin microbiota, from the living environment, and from handling by other individuals. *E. coli* and other facultatively anaerobic species are among the first colonizers of the oxygenous gastrointestinal tract of newborns (Rotimi & Duerden, 1981; Stark & Lee, 1982), and they are commonly transmitted to the infants from the mother by an oral route (Bettelheim *et al.*, 1974; Bettelheim & Lennox-King, 1976; de Muinck *et al.*, 2011). Analysis of fecal samples from over thousand infants at the age of one month showed that the prevalence of *E. coli* is 88% in humans (Penders *et al.*, 2006). During the first months of life the gut progressively becomes more reduced and acifidied, which favours the growth of obligately anaerobic bacteria, mainly *Bifidobacterium* and *Bacteroides*. After a series of successions of bacterial communities, facultatives decline in numbers and the gut microbiota attains its adult-like composition dominated by obligate anaerobes (Palmer *et al.*, 2007). In healthy adults, the numbers of anaerobic bacteria in the intestine are 100-1000 fold over those of aerobic bacteria (Hoogkamp-Korstanje *et al.*, 1979) and the intestinal microbiota is dominated by two bacterial phyla, the Gram-negative Bacteroidetes and the Gram-positive Firmicutes, which make up over 90% of the known phylotypes (Eckburg *et al.*, 2005; Ley *et al.*, 2008; Qin *et al.*, 2010). *E. coli* is able to persist as a resident member of the normal microbiota of the large intestine, especially in the caecum and the colon, from infancy to adulthood. It also frequently colonizes vaginal and endocervical mucosal surfaces (Guiral *et al.*, 2011; Gupta *et al.*, 1998; Krohn *et al.*, 1997; Obata-Yasuoka *et al.*, 2002; Watt

et al., 2003). *E. coli* is considered an important and widespread human commensal (Tenaillon *et al.*, 2010) that benefits the host by inducing colonization resistance to invading pathogens (Freter & Abrams, 1972; Hudault *et al.*, 2001; Leatham *et al.*, 2009; Schamberger *et al.*, 2004; Schultz, 2008).

1.1.2 *E. coli* as a pathogen

E. coli also is significant from the perspective of human and animal health because it is capable of causing a wide range of severe intestinal and extraintestinal infections (Croxen & Finlay, 2010; Kaper *et al.*, 2004). It is believed to kill more than two million humans annually. Indeed, pathogenic *E. coli* is a leading cause of diarrhea-associated deaths worldwide and accounts in developing countries for 30 to 40 percent of acute diarrhea in children under the age of five (O'Ryan *et al.*, 2005). *E. coli* also is the predominant causative agent of Gram-negative neonatal meningitis and responsible for a third of the cases of neonatal meningitis (Gaschignard *et al.*, 2011). Moreover, the species has significant economic impacts as being the most common cause of uncomplicated urinary tract infection (UTI) (Gaynes *et al.*, 2005; Gupta *et al.*, 1999; Warren, 1996).

Pathogenic *E. coli* strains capable of causing disease in healthy humans are classified into eight well-defined pathovars on the basis of the symptoms they cause and their infection route as well as possession of virulence traits (Croxen & Finlay, 2010; Kaper *et al.*, 2004). The six extensively studied diarrheagenic pathovars of *E. coli*, also designated intestinal pathogenic *E. coli* (IPEC), are enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), enteroaggregative *E. coli* (EAEC), and diffusely adherent *E. coli* (DAEC). IPECs are obligate pathogens. In contrast to diarrheagenic isolates, extraintestinal *E. coli* (ExPEC) strains can benignly inhabit the intestine. The properly defined ExPEC pathovars implicated in extraintestinal infections in humans are uropathogenic *E. coli* (UPEC) and neonatal meningitis *E. coli* (NMEC). Besides the above-mentioned pathovars, there are several additional, less characterized *E. coli* pathovars (Kaper *et al.*, 2004), such as Crohn's disease-associated adherent-invasive *E. coli* (AIEC; Boudeau *et al.*, 1999) and avian pathogenic *E. coli* (APEC; Dho-Moulin & Fairbrother, 1999), responsible for colibacillosis in poultry. Most of the *E. coli* pathovars described above are not highly invasive and rarely cause a systemic infection in adult humans.

Pathogenic *E. coli* isolates possess a spectrum of virulence-associated factors that enable them to colonize and cause disease and to manifest distinct features in pathogenesis (Croxen & Finlay, 2010). Such virulence factors include adhesins, invasins, toxins, iron-acquisition systems, and host defense-avoidance mechanisms to attack crucial intracellular pathways in their hosts and to subvert host defences and other important cellular functions.

1.1.3 Neonatal meningitis *E. coli* isolates

NMEC is one of the leading causative agents of neonatal meningitis, an inflammation of the meninges in newborn infants. NMEC and group B *Streptococcus* account for over 80% of cases of neonatal bacterial meningitis in industrialized countries. Despite advances in antibiotics and infant intensive care, the disease continues to be associated with high rates of

mortality and permanent sequelae in survivors worldwide. Mortality is about 10% to 15% in industrialized countries, whereas in developing countries the rate is notably higher, in the range of 40%–58% (Furyk *et al.*, 2011; Gaschignard *et al.*, 2011). NMEC strains are genetically closely related and distributed only in few serotypes (Bonacorsi & Bingen, 2005), and strains possessing the K1 capsule antigen are predominant (80-88%) among isolates from neonatal *E. coli* meningitis (Bingen *et al.*, 1997; Johnson *et al.*, 2002; Korhonen *et al.*, 1985; Robbins *et al.*, 1974). K1 is mostly associated with O18 antigen (Johnson *et al.*, 2002; Korhonen *et al.*, 1985; Sarff *et al.*, 1975), and the common and worldwide distributed O18:K1:H7 clone represents the most studied serotype. NMEC strains are closely related to a subset of APEC and UPEC strains and there is no definite host- or pathovar-specific virulence gene pattern among these groups (Bauchart *et al.*, 2010; Ewers *et al.*, 2007; Moulin-Schouleur *et al.*, 2006; Moulin-Schouleur *et al.*, 2007). Thus, poultry may be a food-borne source of ExPEC strains causing disease in humans, and APEC isolates constitutes a high zoonotic risk and a reservoir for virulence-associated genes of human ExPEC (Tivendale *et al.*, 2010). In adults, *E. coli* K1 strains are often members of the normal microflora (Obata-Yasuoka *et al.*, 2002; Sarff *et al.*, 1975), and they are able to induce meningitis only in exceptional conditions, such as neurosurgery and trauma. However, vertical transmission of NMEC from the microbiota of the maternal genital tract to the infant is considered as a main mode of acquisition of NMEC (Sarff *et al.*, 1975), hence ability to colonize the female genital tract is an aspect of NMEC pathogenesis.

The multiplication in bloodstream is a critical early phase of the K1 infection and a key step in the pathogenesis of the bacterial meningitis. NMEC gains access to the bloodstream by direct hematogenous spread, or by crossing the mucosal barrier and entering the bloodstream after ingestion and colonization of the host intestinal mucosa (Glode *et al.*, 1977; Pluschke *et al.*, 1983; Zelmer *et al.*, 2008). A high degree of bacteraemia ($>10^3$ colony-forming units/ml) is a prerequisite for a penetration of NMEC across the blood-brain barrier (Dietzman *et al.*, 1974; Kim *et al.*, 1992), which is characterized by a microvascular endothelial cells with tight junctions (Rubin & Staddon, 1999). Similar linkage between the magnitude of bacteremia and the crossing the blood–brain barrier has been observed for group B *Streptococcus*, *S. pneumoniae*, and *Haemophilus influenzae* type b (Ferrieri *et al.*, 1980; Moxon & Ostrow, 1977; Sullivan *et al.*, 1982). The development of *E. coli* K1 meningitis correlates with the age of the host, and the greatest risk of acquiring meningitis is in the neonatal period (<28 days) (Korhonen *et al.*, 1985; Unhanand *et al.*, 1993). Neonates have deficiencies in each of the major arms of the immune system, including cellular, humoral and innate immunity. It has been suggested that the age dependency of *E. coli* meningitis results from the ability of *E. coli* K1 to evade the immature host defenses to achieve the threshold level of bacteraemia rather than differences between neonates and adults in host tissue interactions with *E. coli* (Stins *et al.*, 1999). After binding to and transmigration through brain microvascular endothelial cells in an enclosed vacuole without intracellular multiplication (Prasadarao *et al.*, 1999) into the central nervous system, NMEC multiplies in the cerebrospinal fluid and ceases expression of the K1 capsule (Zelmer *et al.*, 2008). The bacterial growth elicits inflammatory responses from the host, eventually leading to inflammation of the meninges and blood–brain barrier dysfunction (Kim, 2003).

Several bacterial factors have been recognized to contribute to the development of *E. coli* meningitis in neonates. One of the most important virulence factor is the K1 capsule, a linear homopolymer of α-2,8-linked polysialic acid that is structurally identical to the meningococcal B capsular polysaccharide (Troy, 1992). The K1 capsule is protective antigen of NMEC that confers serum resistance and an antiphagocytic barrier to the bacterium (Allen *et al.*, 1987; Bortolussi *et al.*, 1983; Cross *et al.*, 1986; Mittal & Prasadarao, 2008). It also prevents lysosomal fusion in microvascular endothelial cells when crossing the blood-brain barrier (Kim *et al.*, 2003). Notable, Finne (1982) reported the presence of large amounts of α-2,8-linked polysialic acid in developing mammalian brain tissue, and the expression was shown to be associated predominantly with neural cell adhesion molecules (NCAM; Finne *et al.*, 1983; Hoffman *et al.*, 1982). Thus, neonatal immune responses to *E. coli* K1 are likely to be restricted due to molecular mimicry of the polysialic acid capsule to polysialic acid epitope of NCAM on fetal neuronal tissue (Troy, 1992). Efficient replication in bloodstream also demands iron acquisition systems, such as IroN and aerobactin (Khan & Isaacson, 2002; Nègre *et al.*, 2004), and surface expression of OmpA which binds C4b-binding protein (C4bp) and thereby inhibits classical complement pathway (Wooster *et al.*, 2006). OmpA also binds to brain vascular endothelium (Prasadarao *et al.*, 1996; Prasadarao, 2002), induces survival in macrophages (Sukumaran *et al.*, 2004) as well as prevents maturation of dendritic cells (Mittal & Prasadarao, 2008). In addition, multiple adhesion and invasion proteins and protein complexes such as S fimbria, type 1 fimbria, flagellum, type III secretion system, IbeA, and CNF1 have been associated with NMEC penetration into brain microvascular endothelial cells (Huang *et al.*, 1995; Huang *et al.*, 2000; Khan & Isaacson, 2002; Parkkinen *et al.*, 1988b; Parthasarathy *et al.*, 2007; Teng *et al.*, 2005; Yao *et al.*, 2009). Although the bacterial factors mentioned have shed some light on the mechanisms of the development of NMEC infections, the pathophysiology of *E. coli* neonatal meningitis is still rather unclear (Kim, 2003; Kim, 2008), and a more complete understanding of pathogenic determinants of NMEC strains is required for prevention of this complex and multistage disease.

1.1.4 Genome dynamics and evolution of *E. coli*

The recent large outbreak of severe hemolytic uremic syndrome and bloody diarrhoea in Germany caused by Shiga toxin-producing *E. coli* O104:H4, an exceptionally virulent hybrid of EAEC and EHEC (Brzuszkiewicz *et al.*, 2011; Mellmann *et al.*, 2011), highlights the fact that *E. coli* is constantly evolving and able to cause new health threats and mayhem. Behind the phenotypic variety of *E. coli* is an exceptional genomic plasticity, illustrated by the large diversity in genome sizes within the species, even greater than 25% (Bergthorsson & Ochman, 1998; Dobrindt *et al.*, 2003). Since the publication of the first complete *E. coli* genome sequence in 1997 for the nonpathogenic *E. coli* K-12 isolate MG1655 (Blattner *et al.*, 1997), 51 fully sequenced *E. coli* genomes, representing different pathovars and commensals and varying in size from 4.6 to 6.2 Mb, have been released. In addition, 984 currently ongoing projects await completion (http://www.genomesonline.org). Analyses of the genome content and the genome organization of different isolates have revealed that *E. coli* genomes have a mosaic-like structure composed of a conserved part, i.e. the core genes present in all *E. coli* genomes, and a variable part, the so-called flexible gene pool or pan-genome (Dobrindt *et al.*, 2010; Welch *et al.*, 2002). The core genome possesses the genetic information required for essential cellular processes, whereas the pan-genome consists of

e.g. genomic or pathogenicity islands (GEIs, PAIs) and accessory genetic elements (plasmids, transposons, integrons, insertion sequence elements, prophages) and provides additional properties for adaptation to unique environments. The pan-genome of *E. coli* contains approximately 18,000 genes, and a typical genome of an *E. coli* is comprised of about 5,000 genes of which less than 2,000 are conserved in all isolates (Lukjancenko *et al.*, 2010; Rasko *et al.*, 2008; Touchon *et al.*, 2009).

The ability to remodel genetic repertoire by repeated acquiring and losing large chromosomal regions as well as by DNA rearrangements and point mutations has a key impact on the border between commensalism and pathogenicity, i.e. the evolution of pathogenic *E. coli* variants (Ahmed *et al.*, 2008; Dobrindt, 2005; Dobrindt *et al.*, 2010). In *E. coli*, the rate of recombination exceeds the mutation rate by 50-100 folds (Guttman & Dykhuizen, 1994; Touchon *et al.*, 2009). The frequent and widespread recombination within the species drives the clonal divergence. In addition, the *E. coli* genome is particularly permissive to lateral gene transfer, an important mechanism to rapidly exchange genetic information between different bacteria by transformation, transduction, and conjugation, and harbours regulatory systems that assimilate foreign genes into pre-existing regulatory networks (Dobrindt *et al.*, 2010). In a pioneering study, Lawrence and Ochman (1998) estimated that 18% of the genome of the K-12 strain MG1655 consists of laterally transferred genes, acquired in over 200 separate events since its divergence from *Salmonella*. Most of the gene flux takes place at particular regions of high instability, the integration hotspots, and the well-known hotspots of gene acquisition, tRNA genes, counts only a minor fraction of these (Touchon *et al.*, 2009). Virulence factors are frequently encoded by genes located on mobile DNA elements, and therefore the pathogenic potential of *E. coli* is strongly influenced by laterally transferred genes. Current dogma suggests that pathogenic variants can rapidly evolve from their commensal ancestors by lateral transfer of blocks of genes that provide a repertoire of novel traits in a single step. For example, the hallmark of the evolution of *Shigella* is the acquisition a large 210 kb plasmid that encodes primary virulence determinants for entry into epithelial cells, intracellular replication and dissemination from cell to cell (Venkatesan *et al.*, 2001). Strains with the same pathotype, including *Shigella* clones, have arisen independently and repeatedly in several lineages as a result of lateral acquisition of genes (Lacher *et al.*, 2007; Lan *et al.*, 2004; Ogura *et al.*, 2009; Pupo *et al.*, 2000; Reid *et al.*, 2000; Sims & Kim, 2011; Turner *et al.*, 2006; Wirth *et al.*, 2006). In addition to lateral gene transfer, also loss and inactivation of genes shape the genomes of *E. coli* isolates. Nonvirulent strains can originate from pathogenic *E. coli* by reductive evolution and virulence attenuation (Zdziarski *et al.*, 2008). The currently used probiotic fecal strain Nissle 1917 (Mutaflor®) and the probiotic-like asymptomatic bacteriuria (ABU) strain 83972 share common ancestor with the highly virulent UPEC strain CFT073 and exemplify deconstructed pathogens (Vejborg *et al.*, 2010).

Despite the massive flux of genes and the genomic plasticity, the population structure of *E. coli* remains largely clonal (Ochman & Selander, 1984; Selander & Levin, 1980; Selander *et al.*, 1986; Touchon *et al.*, 2009). Based on population genetic studies, *E. coli* strains cluster into four major phylogenetic groups (designated A, B1, D, and B2) along with two minor groups (E and F) (Clermont *et al.*, 2000; Herzer *et al.*, 1990; Tenaillon *et al.*, 2010). Group A and B2 isolates are generally predominant in humans; however socioeconomic, climatic,

and geographic factors have impact on the prevalence of phylogenetic groups among different human populations (Tenaillon *et al.*, 2010; White *et al.*, 2011). Each phylogenetic group contains varying proportions of non-pathogens and different pathogroups. The majority of commensal strains are members of the group A and B1, whereas human ExPEC strains cluster mostly into group B2 and to a lesser extent group D (Bingen *et al.*, 1998; Duriez *et al.*, 2001; Escobar-Páramo *et al.*, 2004; Moulin-Schouleur *et al.*, 2007; Picard *et al.*, 1999). In contrast, diarrheagenic strains are found throughout the phylogeny, though those associated with severe and acute diarrhoea lie outside groups B2 and D (Escobar-Páramo *et al.*, 2004). Several studies suggest that the B2 group is the ancestral phylogenetic group of *E. coli* (Escobar-Páramo *et al.*, 2004; Lecointre *et al.*, 1998; Sims & Kim, 2011). The members of the group exhibit the highest genetic diversity (Touchon *et al.*, 2009) and their genetic background favors acquisition and expression of the laterally acquired genes (Escobar-Páramo *et al.*, 2004). Moreover, the B2 strains are well adapted for asymptomatic human colonization (Moreno *et al.*, 2009; Nowrouzian *et al.*, 2005; Obata-Yasuoka *et al.*, 2002; Watt *et al.*, 2003; Zhang *et al.*, 2002). This indicates that the normal microbiota is an important reservoir for ExPEC strains.

1.2 Fimbriae of *E. coli*

E. coli expresses a multitude of cell surface proteins and their complexes that interact with host and environmental surfaces. The adhesion to surfaces is a prerequisite for bacterial colonization. In microbial pathogenesis, the specific adhesion of the pathogen to host surface receptors is an important determinant in tissue tropism and promotes the pathogen's ability to resist mechanical defences, thereby enabling propagation and the initiation of pathological processes. The functional parts of the adhesins are often localized in peripheral regions of the cell wall (Donnenberg, 2000). Enteric bacteria frequently express their adhesins on thin filamentous protein appendages, referred to as fimbriae (according to Duguid *et al.*, 1955; latin for 'fibre') or pili (according to Brinton, 1959; latin for 'hair'), that extend out from the cell surface, beyond possible capsular and lipopolysaccharide layers, and represent the predominating adhesin type studied in Gram-negative bacteria. Houwink and van Iterson (1950) as well as Duguid *et al.* (1955) demonstrated the presence of fimbriae in *E. coli* as early as the 1950s, and many fimbrial types have subsequently been identified (reviewed in Klemm *et al.*, 2010; Korea *et al.*, 2011), but nevertheless our knowledge of them is still limited and many are poorly characterized, especially the receptors involved in their recognition. All fimbriae do not display adhesive properties; some fimbrial appendages have specialized roles in DNA or protein transport involved in bacterial conjugation or injection of virulence factors into the host cell (Fronzes *et al.*, 2008).

1.2.1 Structure and assembly of fimbriae

A fimbriated bacterial cell generally has a multitude of, up to several hundreds, fimbriae distributed peritrichously over the entire surface of the bacterium (Edwards & Puente, 1998); however, some fimbriae are expressed in a polar orientation (Girón *et al.*, 1991). Because of their abundance on the cell surface, structural fimbrial components are attractive targets for vaccine development and the prevention of bacterial infections. Fimbriae of Gram-negative bacteria are formed by the noncovalent association of hundreds of helical assembled fimbrillins into long filamentous rod-like, or coiled, surface protein polymers, usually 1-2

μm in length and 5–10 nm in diameter (Proft & Baker, 2009). By contrast, the fimbrial subunits of Gram-positive bacteria are covalently linked to each other and to the peptidoglycan cell wall through the formation of intermolecular isopeptide bonds by specific sortases (Hendrickx *et al.*, 2011; Scott & Zähner, 2006). In addition to the repeating major subunit proteins, the fimbrial filament contains a few (or several) minor subunits both in Gram-negative and Gram-positive bacteria. In Gram-negative bacteria the adhesive potential of fimbriae is commonly mediated by the properties of the adhesin molecule located at the distal end of the filament. However, in some fimbrial types the major subunit can play both a structural and an adhesive role (Bakker *et al.*, 1992; Jacobs *et al.*, 1987; Jansson *et al.*, 2006; Van Loy *et al.*, 2002). The adhesion properties also depend on the quaternary structure of the fimbrial filament, particularly when the fimbriated bacterium is exposed to strong hydrodynamic shear forces (Duncan *et al.*, 2005; Forero *et al.*, 2006). Recent force-measuring optical tweezer analyses on a single fimbria level have revealed that the quaternary architecture of helix-like fimbriae permits repeated elongation (5–7 times of the fimbrial length) and contraction of the filament. The flexibility increases the ability of a bacterium to withstand shear forces and allows a redistribution of shear stress over a number of fimbriae (Andersson *et al.*, 2008; Axner *et al.*, 2011).

The adhesive fimbriae of Gram-negative bacteria are classified into four major assembly pathways on the basis of their assembly mechanism: the chaperone-usher (CU) pathway, the extracellular nucleation/precipitation pathway (curli), the alternate chaperone pathway (CS fimbriae) and type IV pathway (Fronzes *et al.*, 2008; Nuccio & Bäumler, 2007; Soto & Hultgren, 1999). These pathways, with the exception of type IV, utilize the general secretion pathway (Sec-pathway) for translocation of fimbrial subunits across the inner membrane. The unfolded subunit proteins are targeted to the Sec-translocase by their cleavable N-terminal signal sequence. In contrast, type IV fimbriae assembly does not involve periplasmic intermediates; the subunits are exported from the cytoplasm to the extracellular environment in one step via an assembly machinery that spans both inner and outer membranes. Of the four classes, the CU fimbriae form the most abundant group of bacterial cell surface filaments and are the most extensively characterized. On the basis of the amino acid sequences of the outer membrane usher protein, the CU family comprises at least 189 members that are divided into six major phylogenetic clades termed α-, β-, γ-, κ-, π-, and σ-fimbriae (Nuccio & Bäumler, 2007). In this pathway, the translocated periplasmic subunits form complexes with specific chaperone proteins that assist in protein folding by completing the Ig-like fold of subunits with its donor β-strand, prevent premature assembly of the subunits, and transport the subunits to an outer-membrane usher. The ordered non-covalent polymerization of subunits occurs at the usher which recruits chaperone–subunit complexes, releases fimbrial subunits from the chaperones, and adds the subunits to the base of the growing fimbria. The usher protein forms a translocation pore in the outer membrane and transports the growing fimbria to the cell surface (Kline *et al.*, 2010; Waksman & Hultgren, 2009). Several small-molecule compounds, termed pilicides, have been identified to prevent CU fimbriae biogenesis process and therefore have potential therapeutic value by blocking adhesive organelles important in colonization. The designed compounds belong to a family of bicyclic 2-pyridones and interfere with the targeting of chaperone–subunit complexes to the usher by binding to the chaperone (Pinkner *et al.*, 2006; Svensson *et al.*, 2001; Åberg & Almqvist, 2007). Similarly, bicyclic 2-pyridones were shown to inhibit extracellular

nucleation/precipitation assembly pathway of adhesive amyloid fimbriae termed curli (Cegelski *et al.*, 2009).

1.2.2 Fimbriae as adhesins

Fimbriae are generally described as adhesive organelles that mediate bacterial attachment and subsequent colonization, pathogen–host interactions, and tropism determination addressing bacteria to a specific site of infection (Table 1). The most studied fimbriae, type 1 fimbria and P fimbria (pyelonephritis-associated fimbria), are composed of a rigid rod and a fibrillar tip (Jones *et al.*, 1995; Kuehn *et al.*, 1992; Lindberg *et al.*, 1987). Type 1 fimbriae are expressed on the surfaces of most *E. coli* isolates as well as on most members of the *Enterobacteriaceae* family, and they mediate binding to D-mannosylated glycolipids and glycoproteins such as uroplakins via the adhesin FimH, which is located at the tip of the fimbria and is also interspersed along the fimbrial filament (Baorto *et al.*, 1997; Hung *et al.*, 2002; Klemm & Christiansen, 1987; Krogfelt *et al.*, 1990; Zhou *et al.*, 2001). The type 1 fimbriae of *E. coli* are important for intestinal colonization and subsequent transmission of *E. coli* in the host population (Bloch *et al.*, 1992) as well as bacterial attachment to the bladder epithelium and persistent colonization of *E. coli* in the lower urinary tract (Connell *et al.*, 1996; Hung *et al.*, 2002; Wright *et al.*, 2007; Zhou *et al.*, 2001). Minor allelic variation of the FimH receptor-binding domain affects tissue tropism through changes in the mannose receptor specificity (Pouttu *et al.*, 1999; Sokurenko *et al.*, 1994; Sokurenko *et al.*, 1995; Sokurenko *et al.*, 1998; Weissman *et al.*, 2007). Similarly, UPEC strains associated with acute pyelonephritis express several variants of the P fimbrial tip adhesin PapG, which differ in their receptor specificity to different Gal(α1-4)Galβ moieties present in the globoseries of membrane glycolipids on erythrocytes of the P blood group and on uroepithelial cells (Leffler & Svanborg-Edén, 1981; Lund *et al.*, 1987; Strömberg *et al.*, 1990). Further, the two other minor proteins (PapE and PapF) at the tip structure mediate bacterial adherence to immobilized fibronectin and to basolateral aspects of renal tubuli (Westerlund *et al.*, 1991), suggesting that P fimbriae are multifunctional adhesive complexes. Similar dual epithelial and extracellular matrix (ECM)-binding is shown by the Dr adhesin of *E. coli* which binds to the decay accelerating factor on epithelium and to the N-terminal 7S fragment of type IV collagen (Nowicki *et al.*, 1988; Nowicki *et al.*, 1993; Nowicki *et al.*, 2001; Westerlund *et al.*, 1989b). Adhesiveness to ECM is of importance in bacterial colonization of damaged tissue sites and has been proposed to potentiate bacterial spread through tissue barriers, e.g. from the intestine into the circulation, to cause systemic infection (Westerlund & Korhonen, 1993).

Table 1. Functions of *E. coli* fimbriae and examples of representative fimbrial types.

1. Adhesion on eukaryotic cells			
Fimbrial type	Host cells	Receptor specificity	Reference
type 1	Various cell types, such as erythrocytes; bladder, kidney, and intestinal epithelial cells; macrophages; brain microvascular endothelial cells	D-mannose oligosaccharides, uroplakins, CD48, leukocyte integrins CD11 and CD18, granulocyte membrane antigen, Tamm-Horsfall glycoprotein	Baorto *et al.*, 1997; Edén & Hansson, 1978; Gbarah *et al.*, 1991; Hung *et al.*, 2002; Isaacson *et al.*, 1978; Parkkinen *et al.*, 1988a; Salit & Gotschlich, 1977; Sauter *et al.*, 1991; Teng *et al.*, 2005; Virkola, 1987; Zhou *et al.*, 2001
P	erythrocytes; kidney, bladder, and intestinal epithelial cells	Gal(α1–4)Galβ, glycolipids GbO3, GbO4, and GbO5	Adlerberth *et al.*, 1995; Hull *et al.*, 1981; Korhonen *et al.*, 1982; Korhonen *et al.*, 1986; Kuehn *et al.*, 1992; Melican *et al.*, 2011; Strömberg *et al.*, 1990
S	erythrocytes; kidney, intestinal and brain epithelial cells; vascular endothelial cells	sialyl(α2–3)Galβ, Tamm-Horsfall glycoprotein	Adlerberth *et al.*, 1995; Korhonen *et al.*, 1984; Korhonen *et al.*, 1986; Moch *et al.*, 1987; Parkkinen *et al.*, 1986; Parkkinen *et al.*, 1988a; Parkkinen *et al.*, 1988b
curli	bladder, kidney, and intestinal epithelial cells	contact phase proteins	Ben Nasr *et al.*, 1996; Kikuchi *et al.*, 2005; Wang *et al.*, 2006
Dr	erythrocytes; bladder, and intestinal epithelial cells; kidney substructures	decay-accelerating factor (DAF, CD55), carcinoembryonic antigen -related cell adhesion molecules (CEACAMs)	Berger *et al.*, 2004; Guignot *et al.*, 2000; Korotkova *et al.*, 2007; Nowicki *et al.*, 1988; Nowicki *et al.*, 1993; Nowicki *et al.*, 2001; Van Loy *et al.*, 2002
F1845	erythrocytes, intestinal epithelial cells	DAF, CEACAMs	Berger *et al.*, 2004; Guignot *et al.*, 2000; Van Loy *et al.*, 2002
CFA/I	erythrocytes, intestinal epithelial cells	sialoglycoprotein, glycosphingolipids	Anantha *et al.*, 2004; Bühler *et al.*, 1991; Jansson *et al.*, 2006; Pieroni *et al.*, 1988
Mat/ECP	larynx, cervical, and intestinal epithelial cells	-	Avelino *et al.*, 2010; Lasaro *et al.*, 2009; Rendón *et al.*, 2007; Saldaña *et al.*, 2009
2. Invasion into eukaryotic cells			
Fimbrial type	Host cells		Reference
type 1	bladder and kidney epithelial cells; brain microvascular endothelial cells		Li *et al.*, 2009; Martinez *et al.*, 2000; Teng *et al.*, 2005
curli	cervical, bladder, and larynx epithelial cells		Gophna *et al.*, 2001; Gophna *et al.*, 2002; Uhlich *et al.*, 2009; Wang *et al.*, 2006

Dr	cervical and ovary epithelial cells	Das *et al.*, 2005; Goluszko *et al.*, 1997
HCP	intestinal epithelial cells	Xicohtencatl-Cortes *et al.*, 2009
3. Binding to ECM proteins		
Fimbrial type	Receptors	Reference
type 1	fibronectin, laminin, type I and type IV collagens	Kukkonen *et al.*, 1993; Pouttu *et al.*, 1999; Sokurenko *et al.*, 1994
P	fibronectin	Westerlund *et al.*, 1989a; Westerlund *et al.*, 1991
S	fibronectin, laminin	Sarén *et al.*, 1999; Virkola *et al.*, 1993
curli	fibronectin, laminin	Gophna *et al.*, 2002; Olsén *et al.*, 1989; Olsén *et al.*, 1993
Dr	type IV collagen	Korotkova *et al.*, 2007; Selvarangan *et al.*, 2004; Westerlund *et al.*, 1989b
G	laminin	Saarela *et al.*, 1996
LPF	fibronectin, laminin, type IV collagen	Farfan *et al.*, 2011
HCP	fibronectin, laminin	Xicohtencatl-Cortes *et al.*, 2009
4. Biofilm formation		
Fimbrial type	Target	Reference
Type 1	polyvinyl chloride (PVC), polystyrene; intestinal and bladder epithelial cells	Bollinger *et al.*, 2006; Hancock *et al.*, 2011; Orndorff *et al.*, 2004; Pratt & Kolter, 1998; Wright *et al.*, 2007
curli	polystyrene, polyurethane, Thermanox™ plastic, stainless steel, glass, Teflon	Kikuchi *et al.*, 2005; Prigent-Combaret *et al.*, 2000; Uhlich *et al.*, 2009; Vidal *et al.*, 1998
F1C	polystyrene, glass	Hancock *et al.*, 2011; Lasaro *et al.*, 2009; Xicohtencatl-Cortes *et al.*, 2009
HCP	glass	Xicohtencatl-Cortes *et al.*, 2009
AAF/II	polystyrene, glass	Sheikh *et al.*, 2001
5. Protection against antimicrobials		
Fimbrial type	Target	Reference
curli	peptide LL-37	Kai-Larsen *et al.*, 2010
longus	lysozyme and several antibiotics	Clavijo *et al.*, 2010
6. Plasminogen activation by acting as a plasminogen receptor		
Fimbrial type		Reference
type 1		Parkkinen & Korhonen, 1989
P		Parkkinen & Korhonen, 1989
S		Parkkinen *et al.*, 1991
curli		Sjöbring *et al.*, 1994
G		Kukkonen *et al.*, 1998

7. Twitching motility	
Fimbrial type	Reference
longus	Mazariego-Espinosa *et al.*, 2010
HCP	Xicohtencatl-Cortes *et al.*, 2009

The adhesins of both type 1 and P fimbriae are of the catch-bond type in which the receptor-ligand interaction is enhanced by mechanical force pulling the ligand and its receptor apart (Le Trong *et al.*, 2010; Nilsson *et al.*, 2006; Thomas *et al.*, 2002; Thomas *et al.*, 2004). Although the helical filaments of type 1 and P fimbriae share structural similarities (Hahn *et al.*, 2002), the characterization of their intrinsic biomechanical properties have uncovered differences, e.g. type 1 fimbria was shown to be more rigid and responds faster to an external force. It has been suggested that the dissimilarities have evolved to optimize functions of the fimbriae to the environmental conditions in different parts of the urinary tract; type 1-fimbriated bacterial cells attached to epithelial cells in the urethra can resists high forces for a shorter time during the irregular urine flow, while UPEC ascending to kidney utilizes P fimbriae for adhesion in the upper urinary tract where a more constant urine flow is present (Andersson *et al.*, 2006; Andersson *et al.*, 2007).

Fimbriae also participate in other functions than host cell adhesion, such as biofilm formation, invasion, plasminogen activation, and twitching motility (Table 1; Kline *et al.*, 2010; Proft & Baker, 2009). For example, several types of fimbriae can immobilize a circulating host protease precursor plasminogen on the bacterial surface and thus enhance conversion of plasminogen into active plasmin by mammalian plasminogen activators. Adherence to ECM and creation of bacterial surface proteolysis with the aid of fimbrial structures can enhance bacterial invasion through tissue barriers and thus promote bacterial migration in the body (Lähteenmäki *et al.*, 2001).

1.2.3 Fimbriae as biofilm-promoting factors

Biofilms are social consortia of cells that are surface-bound or localized to an air–liquid interface and encased in a self-produced extracellular matrix composed of highly hydrated exopolysaccharides, proteins, nucleic acids, and lipids (Flemming & Wingender, 2010). This communal lifestyle is an integral growth mode for most bacterial species; it allows bacteria to survive in stressful conditions and promotes transmission to new niches. Residence in a biofilm offers many advantages that are not equally enjoyed by planktonic bacteria, for example, improved acquisition of nutrients and transmissible genetic elements, and enhanced protection against antimicrobial agents, host-defence mechanisms, protozoan grazers, desiccation, hydrodynamic shear forces, ultraviolet radiation, and other environmental assaults (Hall-Stoodley et al., 2004; Karatan & Watnick, 2009). Biofilm formation occurs on virtually any surface exposed to bacterial populations; notably it was discovered that UPEC can form biofilm-like communities even in intracellular niche within the bladder superficial umbrella cells (Anderson *et al.*, 2003) via a process dependent on type 1 fimbriae (Wright *et al.*, 2007).

Transition from single planktonic cells to complex three-dimensional biofilm structures occurs through several stages (Fig. 1). This complex process can be viewed as a developmental system (Monds & O'Toole, 2009; O'Toole *et al.*, 2000; Stoodley *et al.*, 2002). The first stage in biofilm formation, a reversible and initial attachment onto a surface, often involves flagella-mediated motility that permits the *E. coli* bacteria to overcome the repulsive forces generated between the two surfaces (Pratt & Kolter, 1998; Wood *et al.*, 2006). The irreversible attachment of cells to the surface necessitates the production of other surface structures, including type 1 fimbriae and curli, to strengthen the adhesion (Ghigo, 2001; Pratt & Kolter, 1998; Reisner *et al.*, 2003; Vidal *et al.*, 1998). An efficient adhesin, such as curli, F conjugative pilus or aggregative adherence fimbriae (AAF), can bypass the need for flagellar motility in the early stages of *E. coli* biofilm development (Prigent-Combaret *et al.*, 2000; Reisner *et al.*, 2003; Sheikh *et al.*, 2001). Next, cell division leads to building of microcolonies, and self-produced extracellular polymeric substances such as colanic acid immobilize biofilm cells to the surface and to each other. As biofilm development proceeds, bacteria undergo profound physiological and metabolic adjustments, including induction or downregulation of fimbriae genes (Domka *et al.*, 2007; Hancock & Klemm, 2007; Schembri *et al.*, 2003). Thus, biofilm cells display different phenotypes relative to planktonic cells (Stewart & Franklin, 2008). Integration of a variety of environmental, cell-to-cell, and intracellular signals into an ordered interplay of regulatory networks determines the architecture of mature biofilms, such as mushroom- or tower-like structures (Karatan & Watnick, 2009; Prüß *et al.*, 2006). Finally, seeding and spread of new biofilms is initiated through detachment of cells and erosion of the mature biofilm as well as by programmed generation and release of differentiated, highly motile cells known as dispersal cells (McDougald *et al.*, 2012).

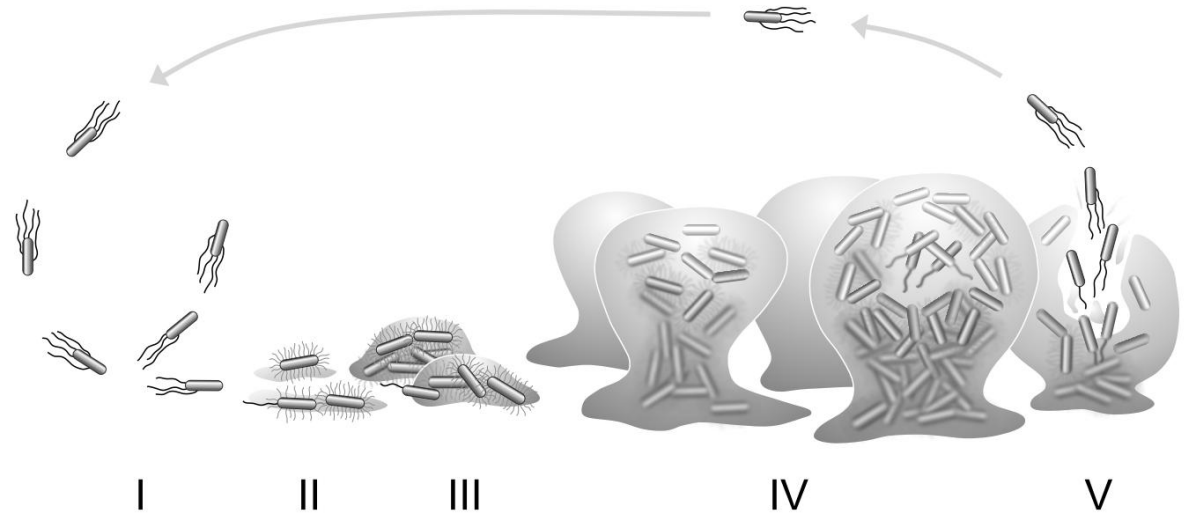

Figure 1. Developmental model of biofilm formation. Biofilm formation starts with the reversible attachment (I) of flagellated bacteria to a surface and proceeds to irreversible attachment (II) via lost of motility and production of specific adhesins such as fimbriae. After stable association with the surface, bacteria form microcolonies and produce an extracellular matrix that holds the cells together (III). Biofilm formation progress to maturation (IV) with characteristic three-dimensional architecture, in some cases pillars with water channels. At the dispersion stage (V), a subpopulation of cells will undergo lysis, and another subpopulation of differentiated, motile cells disperses from the biofilm. Modified from Stoodley *et al.* (2002).

1.3 Genetic organization and regulation of fimbriae in *E. coli*

1.3.1 Fimbrial operons and gene variation

With the availability of complete genome sequences for various *E. coli* isolates, it has become clear that each *E. coli* chromosome harbors large number (≥10), and mainly unique repertoire, of fimbrial operons (Archer *et al.*, 2011; Chaudhuri *et al.*, 2010; Low *et al.*, 2006; Vejborg *et al.*, 2011). For example, the K-12 strain MG1655 harbours 11 operons that belong to chaperone–usher family (Korea *et al.*, 2010). In addition, several *E. coli* plasmids contain gene clusters for fimbriae expression (Czeczulin *et al.*, 1997; Evans *et al.*, 1975; Shipley *et al.*, 1978; Stone *et al.*, 1996). Such diversity in fimbrial systems enables a bacterium to express several types of fimbrial functions and consequently bind to different receptors on different hosts and on different sites within the same host. The apparent functional redundancy also promotes immune evasion. The expression of fimbrial operons are known to be interconnected to limit coexpression of similar fimbrial operons and to prevent excessive immune stimulation, as well as to achieve sequential expression of fimbriae (Holden & Gally, 2004; Totsika *et al.*, 2008). Thus, a single bacterium seldom produces various types of fimbriae simultaneously.

Fimbrial operons share common genetic organization within a certain fimbrial assembly pathway. For example, the chaperone-usher fimbrial operon typically contains genes encoding a major subunit, a chaperone, an usher, a minor subunit, and a tip adhesin. One or two genes encoding fimbriae regulatory proteins are located upstream of the operon (Korea *et al.*, 2011; Nuccio & Bäumler, 2007). In some cases, additional regulatory genes are located downstream of the main fimbrial operon (Sjöström *et al.*, 2009a; Sjöström *et al.*, 2009b). Fimbrial genes display significant sequence divergence, particularly the fimbrial adhesins and the regions of major structural subunit that are predicted to encode surface exposed epitopes (Bauchart *et al.*, 2010; Boyd & Hartl, 1998; Peek *et al.*, 2001; Weissman *et al.*, 2006). As in the case of the variation of the repertoire of fimbrial gene clustres among *E. coli* isolates, the allelic variation increases avoidance of the host immune defense systems and contributes to differences in the receptor-binding specificity. The FimH diversification represents a well-characterized example of pathoadaptation; certain FimH variants in UPEC strains enhance mannose-sensitive binding to host glycoproteins and/or colonization of the urinary tract (Sokurenko *et al.*, 1998; Weissman *et al.*, 2007), while a naturally occurring A62S substitution associated within NMEC strains enables *E. coli* to adhere to collagens (Pouttu *et al.*, 1999).

1.3.2 Regulators controlling fimbriae expression

The regulation of bacterial transcription has been viewed as a rather simple process compared to that in eukaryotes. However, until recently, the technological breakthroughs, high-throughput analyses, and increased sequencing capacity has re-shaped our understanding of the bacterial transcriptome in terms of complexity and versatility. Processes that alter the rates of transcriptional elongation and termination, mRNA degradation, and mRNA translation are no longer seen as curiosities (Belasco, 2010; Serganov & Patel, 2007; Sorek & Cossart, 2010). In bacteria, transcription initiation is a prominent regulated step in gene expression. The *E. coli* K-12 genome contains more than 2400 experimentally

determined transcriptional start sites and over 1200 predicted promoters (Mendoza-Vargas *et al.*, 2009). About 270 species of DNA-binding transcription factors, 20-30 species of RNA polymerase-binding regulatory proteins, and seven species of promoter recognition sigma factors modulate the promoter selectivity of RNA polymerase (Ishihama, 2010). Moreover, a wide variety of post-transcriptional and translational regulatory mechanisms including numerous RNA regulators ensures appropriate protein synthesis (Storz *et al.*, 2011; Waters & Storz, 2009).

The complex interactions between gene repertoire and the regulatory apparatus form multilayered, hierarchical regulatory networks that constantly modulate the cell responses according to the environmental conditions. Two-component signal transduction (TCST) is the predominant bacterial system for sensing environmental factors and transmission of the information to cell interior, and TCST typically is composed of a membrane-associated histidine sensor kinase and a cytoplasmic response regulator (Gao & Stock, 2009). Over 10% of transcription factors in *E. coli* are involved in TCST systems (Oshima *et al.*, 2002; Yamamoto *et al.*, 2005). A specific environmental input signal is perceived by the sensor kinase, which auto-phosphorylates at its transmitter domain and donates the phosphate to the receiver domain of a response regulator, thereby modulating its activity. For instance, cell-surface and cell-cell interactions trigger alterations in the integrity of the bacterial surface, such as membrane perturbations. The alterations of the bacterial envelope are sensed by at least six envelope stress response pathways in *E. coli* (Bury-Moné *et al.*, 2009; McBroom & Kuehn, 2007; Rowley *et al.*, 2006), including Cpx and Bae two-component systems and Rcs phosphorelay system. The latter system is an expanded form of the canonical TCST that involves additional phosphotransfer protein between sensor kinase and response regulator (Clarke, 2010). The Rcs phosphorelay plays a central role in the transition from a motile to a sessile biofilm lifestyle (Huang *et al.*, 2006; Majdalani & Gottesman, 2005) ensuring the timely production of biofilm-related factors, such as flagellum, curli, and type 1 fimbria (Ferrières & Clarke, 2003; Francez-Charlot *et al.*, 2003; Schwan *et al.*, 2007; Vianney *et al.*, 2005).

The assembly of hundreds of fimbriae on cell surface represents a considerable energy sink for the bacterial cell, and may stimulate host innate inflammatory responses during infection. Therefore, *E. coli* have developed a spectrum of sophisticated and finely controlled pathways to modulate fimbriation status (Clegg *et al.*, 2011). Rapidly sensing, interconnected regulatory networks control fimbriae gene expression in response to varying environmental conditions, such as shifts in temperature, pH, osmolarity, and nutrient levels. *E. coli* uses a combination of environmental signals to recognize the microenvironment and to optimize fimbriation in a proper niche at the correct time. Indeed, many of the identified *E. coli* fimbrial operons are poorly expressed under standard laboratory growth conditions (Korea *et al.*, 2010; Low *et al.*, 2006). Generally, the expression is optimal at a temperature of 37°C, the internal temperature of mammalian hosts. For instance, type 1 fimbria (Gally *et al.*, 1993), P fimbria (Göransson & Uhlin, 1984), and S fimbria (Schmoll *et al.*, 1990) have a maximal expression level at 37°C while curli is efficiently transcribed *in vitro* at lower temperatures (Olsén *et al.*, 1989). The production of fimbriae is also dependent on the bacterial growth rate: the assembly of type 1 fimbriae into surface organelles occurs at early logarithmic phase and decreases dramatically as the culture enter the late logarithmic phase

(Dodd & Eisenstein, 1984), whereas curli genes are maximally expressed during stationary phase (Olsén *et al.*, 1993).

Fimbriae promoters are recognized to be under the control of multiple transcription factors. The factors mediating transcriptional regulation of fimbriae expression are either global regulators controlling the expression of a variety of genes throughout the genome or more restricted local regulators encoded by the respective fimbrial operons. The highly expressed histone-like nucleoid structuring protein H-NS is involved in global modulation of gene expression and DNA condensation and preferentially interacts with intrinsically curved DNA, which is commonly found at promoters (Atlung & Ingmer, 1997; Dorman, 2004). H-NS has a central role in the adaptability of cells to changing environmental conditions (Hommais *et al.*, 2001), and it mediates temperature regulation as a repressor in many fimbrial operons, such as the type 1 fimbria (Kawula & Orndorff, 1991), P fimbria (Göransson *et al.*, 1990), S fimbria (Morschhäuser *et al.*, 1993), and curli (Olsén *et al.*, 1993). The relief of H-NS-mediated repression can be achieved by various mechanisms (Stoebel *et al.*, 2008). Most commonly, promoter-specific transcriptional activators act as an antagonist that compete with H-NS for binding to promoter or displace H-NS-nucleoprotein complex from the promoter. For example, expression of the two adjacent divergently transcribed curli operons *csgBA* and *csgDEFG* are repressed by H-NS, and the curli activator CsgD is required for activation of H-NS-silenced curli genes (Hammar *et al.*, 1995). The transcription of *csgD* is under the control of two sigma factors and more than 10 transcription factors, and thus the *csgD* promoter represent one of the most complex promoters in *E. coli* (Ishihama, 2010; Ogasawara *et al.*, 2011).

A cAMP receptor protein CRP is another global regulator involved in gene expression of multiple chaperone-usher fimbrial operons (Båga *et al.*, 1985; Edwards & Schifferli, 1997; Korea *et al.*, 2010; Müller *et al.*, 2009; Weyand *et al.*, 2001). The CRP protein is activated by the binding of a signal metabolite 3′,5′-cyclic adenosine monophosphate (cAMP; Kim *et al.*, 1992; Passner & Steitz, 1997) in response to several environmental conditions such as glucose starvation (Ishizuka *et al.*, 1993). CRP-cAMP directly controls the expression of several hundreds genes including operons needed for carbon source catabolism of non-glucose sugars (Gosset *et al.*, 2004; Grainger *et al.*, 2005; Shimada *et al.*, 2011; Zheng *et al.*, 2004). In contrast to H-NS, CRP-cAMP can both repress and activate fimbriae expression. For example, the expression of type 1 fimbriae is repressed by CRP-cAMP (Müller *et al.*, 2009) while the opposite action is seen for P fimbriae expression (Weyand *et al.*, 2001).

1.3.3 Phase variation

Heterogeneity in surface antigens is a key feature in adaptation of *E. coli* to different niches. In addition to differences in adhesin clusters present in *E. coli* genomes leading to interpopulation diversity, the species can also exhibit cell-to-cell phenotypic variability within a genetically clonal population. Phase variation is an important adaptive strategy to reversibly switch on and off the expression, creating variant subpopulations. Two general processes that give rise to fimbrial phase variation are targeted genetic rearrangement and epigenetic control. Type 1 fimbriae expression is a well-studied example of ON/OFF phase variation due to high-frequency site-specific recombination at specific loci. The orientation

of the invertible DNA element known as the *fim* switch (*fimS*), which is flanked by two 9-bp inverted repeat sequences and contains the promoter for *fim* structural gene expression, determines when the transcription of the *fim* operon occurs (Abraham *et al.*, 1985; Freitag *et al.*, 1985). The precise inversion is catalyzed by the fimbria-specific recombinases FimB and FimE that use the inverted repeats as a substrate (Gally *et al.*, 1996; Klemm, 1986; McClain *et al.*, 1991). The adjustment of the balance between the two fimbriation states responds to environmental signals (Gally *et al.*, 1993; Schwan *et al.*, 2002), and several regulators, such as H-NS, Lrp, IHF, CRP-cAMP, and RcsB, are known to affect the recombination process by directly interacting with the *fimS* and/or modulating the expression of the recombinases (Blomfield *et al.*, 1993; Corcoran & Dorman, 2009; Eisenstein *et al.*, 1987; Gally *et al.*, 1994; Kawula & Orndorff, 1991; Schwan *et al.*, 2007). In addition, sequence polymorphism within and adjacent to the *fimS* results in variation in the switching behaviour among UPEC isolates (Leathart & Gally, 1998).

Phase variation in the expression of P, S, $F165_1$, $F165_2$, F1845, CS31A, and K88 fimbriae does not involve rearrangements in DNA sequences (Blyn *et al.*, 1990; Daigle *et al.*, 2000; Huisman & de Graaf, 1995; Martin, 1996; van der Woude & Low, 1994). Their cell-to-cell variation has an epigenetic basis; the expression is dependent on reversible changes in Dam methylation status of GATC sites located in the promoter region. The mechanism involves deoxyadenine methylase (Dam) and leucine-responsive regulatory protein (Lrp) that compete for access to GATC sites. The resulting changes in methylation in turn alter the binding of transcription factors to the promoter and hence control the initiation of transcription (for review, see Casadesús & Low, 2006).

1.3.4 Post-transcriptional regulation

Post-transcriptional regulation of RNA is a common and efficient strategy for controlling gene expression. Contrary to the general view, it is not a characteristic restricted to eukaryotic mRNA but also plays an important and often underestimated means of modulating bacterial gene expression. mRNA stability is a tightly regulated and integral part of gene expression. Another level of post-transcriptional control takes place during translation; initiation of translation is rate-limiting for protein synthesis and is the main target for translational control. For instance, the expression of the curli regulator CsgD is post-transcriptionally controlled by two antisense small RNAs, OmrA and OmrB, that affect translational efficiency by interacting with the *csgD* 5′-UTR (Holmqvist *et al.*, 2010). *trans*-acting ribonucleases (RNases), RNA binding proteins, and small RNAs continuously intervene mRNA turnover and maturation to adjust mRNA concentrations and to ensure fidelity of the synthesized transcripts (Arraiano *et al.*, 2010; Storz *et al.*, 2011). RNA degradation is also needed to recycle ribonucleotides for incorporation into new RNA molecules. RNase E is a principal endoribonuclease in *E. coli* and acts as a scaffold for a RNA degradosome, a large multiprotein complex involved in degradation of bulk RNA (Carpousis, 2007; Py *et al.*, 1994). In addition, RNase E participates in the processing and maturation of a variety of RNA molecules by cleaving its substrates in A+U-rich regions (Callaghan *et al.*, 2005; McDowall *et al.*, 1994). Fimbrial genes encoding the structural components are generally clustered into transcriptional unit and co-transcribed from a single promoter. The polycistronic operon structure ensures coordination of expression as all

proteins needed for the biogenesis are synthesized concurrently. However, the fimbrial proteins are required in nonstoichiometric amounts; the need for the major fimbrillin is much larger than for the minor subunits and assembly proteins. mRNA processing iniated by endoribonucleolytic cleavage is important mechanism to selectively degrade discrete segments of polycistronic transcripts and to yield transcripts of different stability, thereby achieving differential expression of multiple translational units in polycistronic mRNAs. Such posttranscriptional regulatory process is known to occur in the expression of P- (Bricker & Belasco, 1999; Båga *et al.*, 1988; Nilsson & Uhlin, 1991; Nilsson *et al.*, 1996), S- (Balsalobre *et al.*, 2003), CFA/I- (Jordi *et al.*, 1993), and F1845-fimbriae (Bilge *et al.*, 1993; Koo *et al.*, 2004). In the case of P and S fimbriae expression, the processing involves RNase E-dependent endoribonucleolytic cleavages to achieve a stable transcript encoding the major structural subunit while endoribonucleolytic cleavage of polycistronic mRNA in the operon encoding F1845 fimbriae is independent of RNase E but requires HrpA, a RNA helicase from the DEAH-box family.

Partial termination of transcription also regulates the relative production of the fimbriae proteins (Balsalobre *et al.*, 2003; Båga *et al.*, 1985; Hammar *et al.*, 1995; Jordi *et al.*, 1993). The termination site directly downstream from the major subunit genes dictates whether or not transcription of the downstream genes will occur. An intrinsic, Rho-independent transcriptional terminator has a GC-rich palindromic stem with tetraloop followed by a short uracil-rich stretch of sequence (d'Aubenton Carafa *et al.*, 1990). The termination occurs at or near the U-stretch which form a particularly unstable RNA-DNA hybrid (Martin & Tinoco, 1980), thereby causing pausing of the transcription complex, folding of the terminator hairpin, and finally dissociation of the transcription complex and release of the nascent RNA transcript (Gusarov & Nudler, 1999; Wilson & von Hippel, 1995). Also binding of the regulatory protein Rho to untranslated nascent transcript destabilizes the RNA-DNA hybrid and dissociates the mRNA from RNA polymerase (Richardson *et al.*, 1975; Roberts, 1969). The phase-variable termination of the type 1 fimbriae *fimE* mRNA has been demonstrated to be post-transcriptionally controlled by a Rho-dependent terminator located in the *fimS* switch that terminates transcription when *fimS* is in the OFF orientation (Hinde *et al.*, 2005; Joyce & Dorman, 2002). Moreover, translational control, e.g. through different ribosomal binding sites in polycistronic mRNA, contributes to the proper stoichiometric expression of different genes in polycistronic operons (Deana & Belasco, 2005).

1.3.5 Cross-talk between different fimbrial operons

Multiple studies utilizing mainly UPEC strains have demonstrated that regulatory cross-talk exists between different fimbrial gene clusters (Holden *et al.*, 2001; Holden *et al.*, 2006; Lindberg *et al.*, 2008; Sjöström *et al.*, 2009a; Snyder *et al.*, 2005; Totsika *et al.*, 2008; Xia *et al.*, 2000). Although *E. coli* isolates commonly have the potential to express an array of fimbrial structures, they seldom produce more than one type of fimbria simultaneously on the bacterial surface (Nowicki *et al.*, 1984; Xia *et al.*, 2000). This indicates the precence of regulatory mechanisms that prevent fimbrial coexpression and could promote sequential expression at the single-cell level (reviewed by Holden & Gally, 2004). There can be many copies of *pap* operons encoding P fimbriae within the same chromosome of UPEC isolate. Totsika and co-workers (2008) found that sequence variation in the *pap* regulatory regions as

well as in the P fimbriae regulators PapI and PapB prevent cross-activation of these homologous fimbrial operons. Xia *et al.* (2000) showed that introduction of PapB into *E. coli* K-12 induces on-to-off phase transition frequency of the type 1 fimbriae by blocking FimB recombinase activity at the *fimS* switch and increasing FimE expression. Further studies demonstrated that some PapB paralogues, namely SfaB (S fimbriae), FocB (F1C fimbriae), and PefB (*Salmonella enterica* serovar Typhimurium –plasmid-encoded fimbriae), also affect the expression of type 1 fimbriae in an inverse manner (Holden *et al.*, 2001; Lindberg *et al.*, 2008). The distal end of the main sfa_{II} fimbrial operon of the NMEC strain IHE 3034 contains the $sfaX_{II}$ locus (Sjöström *et al.*, 2009b). The ectopic overexpression of $SfaX_{II}$ down-regulates type 1 fimbriae expression by altering the phase variation control and transcriptional activity from the *fimA* promoter (Sjöström *et al.*, 2009a). Regulation also occurs in the opposite direction; the expression of type 1 fimbriae repress P fimbria expression (Snyder *et al.*, 2005), and PapB has the ability to repress the expression of F1C fimbriae (Lindberg *et al.*, 2008). Moreover, FocB was shown to activate P fimbria expression (Lindberg *et al.*, 2008), representing an example of positive cross-talk between related gene clusters.

1.3.6 Stick or swim: cross-talk between fimbrial operons and flagella system

E. coli is generally motile, and expression of flagella has been shown to contribute to pathogenicity of *E. coli* (Lane *et al.*, 2005; Lane *et al.*, 2007b; Wright *et al.*, 2005). In addition to motility, flagella filaments of *E. coli* mediate adherence to host cells and mucus (Erdem *et al.*, 2007; Girón *et al.*, 2002; Mahajan *et al.*, 2009; Parthasarathy *et al.*, 2007; Roy *et al.*, 2009), promote host cell invasion (Luck *et al.*, 2006; Parthasarathy *et al.*, 2007; Pichon *et al.*, 2009), and contribute to biofilm formation (Pratt & Kolter, 1998; Wood *et al.*, 2006). The biosynthesis and operation of the large flagella organelles necessitate considerable investments of energy and resources. Over 60 genes are required for the process of assembly and function of the flagellum (Brown *et al.*, 2009). In addition, flagellation may provoke innate immune system during infection (Hayashi *et al.*, 2001). Hence, tight and precise regulation of flagella regulon in response to environmental changes is necessary. As a result, the flagella genes are under a complex regulatory cascade of three classes and their expression is controlled in a hierarchical manner to facilitate ordered secretion and assembly process (Kalir *et al.*, 2001; Soutourina & Bertin, 2003). The heteromultimeric master regulator $FlhD_4C_2$ is situated at the top of the hierarchy and dictates when flagella are produced (Komeda, 1986; Kutsukake *et al.*, 1990; Liu & Matsumura, 1994). Expression of *flhDC* operon is the main regulatory point of flagella system. The *flhDC* promoter is controlled by many environmental factors and several regulators including H-NS (Bertin *et al.*, 1994; Soutourina *et al.*, 1999), CRP-cAMP (Silverman & Simon, 1974; Soutourina *et al.*, 1999), OmpR (Shin & Park, 1995), and the response regulator RcsB with its accessory cofactor RcsA (Francez-Charlot *et al.*, 2003). After the initiation of *flhDC* transcription, the stability of formed message is regulated by the small RNA-binding protein CsrA (Wei *et al.*, 2001). $FlhD_4C_2$ is further controlled at the level of protein stability by the protease ClpXP (Tomoyasu *et al.*, 2003).

To avoid simultaneous expression of attachment-promoting fimbriae and motility-enhancing flagella, bacteria have evolved mechanisms for reciprocal regulation of these traits (Akerley

et al., 1995; Clegg & Hughes, 2002; Gardel & Mekalanos, 1996; Holden & Gally, 2004; Li *et al.*, 2001; Ogasawara *et al.*, 2011; Pesavento *et al.*, 2008; Saini *et al.*, 2010; Simm *et al.*, 2004). Indeed, the production of flagella in *E. coli* is decreased after transition from a planktonic to an adhesive lifestyle (Cleary *et al.*, 2004; Prigent-Combaret *et al.*, 1999). Lane *et al.* (2007a) demonstrated that abundance of type 1 fimbrial structures at the cell surface decreases motility of UPEC. Moreover, overexpression of P fimbriae protein PapX in UPEC or S fimbriae protein $SfaX_{II}$ in NMEC had a repressing effect on motility (Simms & Mobley, 2008; Sjöström *et al.*, 2009a). Similarly, motility of the uropathogen *Proteus mirabilis* was affected upon elevated expression of *mrpJ*, the last gene of the *mrp* operon encoding for the *Proteus*-like MR/P fimbriae (Li *et al.*, 2001; Pearson & Mobley, 2008). It has been shown that both PapX and MrpJ directly bind to the *flhDC* flagellar master operon, leading to a repression of transcription of the master regulator $FlhD_4C_2$ (Pearson & Mobley, 2008; Reiss & Mobley, 2011). SfaX has 96% amino acid identity with PapX (Sjöström *et al.*, 2009b) and thus likely has affinity towards *flhDC* promoter; however, this has not yet been established.

1.4 The common and conserved *mat* fimbrial operon

Mat fimbria was discovered in Timo Korhonen's laboratory during studies on the S- and type 1-fimbria-negative derivative of NMEC isolate IHE 3034. The expression of the Mat fimbria *in vitro* was found to be restricted to the genetically conserved 18ac:K1:H7 clonal group of NMEC strains cultivated at a low temperature (20°C); hence the peritrichous filament structures were named meningitis associated and temperature regulated (Mat) fimbriae. Characterization of the fimbria revealed that it has a diameter of 5 to 7 nm and a length varying between 0.4 to 5 µm, and the filament is composed of a 18 kDa fimbrillin protein encoded by *matB* gene (Pouttu *et al.*, 2001; see Fig. 2). The filaments protruding from the bacterial surface has been observed to coil together (Fig. 2) forming 12-nm-wide structures that appear to tether bacterial cells to each other (Avelino *et al.*, 2010; Garnett *et al.*, 2012; Rendón *et al.*, 2007). In addition to *matB* gene, *matA* and *matC* were shown to

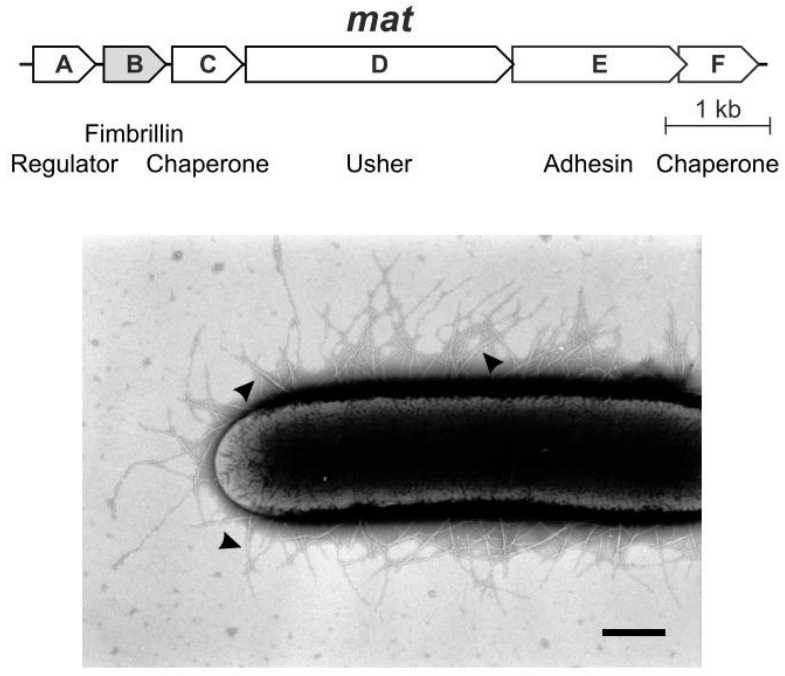

Figure 2. Production of Mat fimbriae in *E. coli*. Schematic presentation of the *mat* operon in the chromosome of *E. coli* IHE 3034, and the predicted function of the gene products according to Pouttu *et al.* (2001), Rendón *et al.* (2007), and Garnett *et al.* (2012). The electron micrograph shows the negatively stained *E. coli* K-12 strain HB101 (pMAT6) cell carrying the *matBCDEF* genes from IHE 3034 under the inducible *ptrc* promoter. Arrowheads point to representative intertwined Mat fimbriae. No fimbrial structures were seen in HB101 harbouring the pSE380 as a vector control. Size bar, 0.5 µm.

affect Mat fimbria biogenesis (Pouttu *et al.*, 2001). These genes are a part of compact 7-kb gene cluster located on the chromosome of IHE 3034 (Fig. 2). Based on usher protein comparison, the Mat fimbria belongs to the alternate chaperone–usher family (α-fimbriae), a relatively small but highly divergent group of fimbriae (Nuccio & Bäumler, 2007). However, the whole 7-kb *mat* gene cluster region does not share significant homology with DNA encoding other fimbrial adhesins (Pouttu *et al.*, 2001). Interestingly, the *mat* region is highly similar, 97.8 % identical in sequence, to the corresponding region in the MG1655 (K-12) chromosome; however, MG1655 is incapable of Mat fimbriae expression in standard growth medium (Pouttu *et al.*, 2001). A considerable amount of evidence suggests that MatB fimbrillin represents a very common and highly conserved surface structure in *E. coli*; over 91% of the *E. coli* isolates representing different pathovars as well as commensals harbor *matB* and nearly 60% express the fimbria (Table 2). Because of the wide distribution of the fimbrillin gene among the *E. coli*, Rendón and co-workers (2007) proposed an alternative name *E. coli* common pilus (ECP) for the fimbria.

Table 2. Distribution of *matB* gene and surface expression of Mat fimbriae in human *E. coli* isolates.

Pathogroup[a]	Presence of *matB*	Mat expression	Reference
Various (mainly NMEC and UPEC)	25/27 (91.7%)	5/27 (18.5%)[b]	Pouttu *et al.*, 2001
Various (mainly IPEC and NFEC)	169/176 (96.0%)	121/169 (71.6%)[c]	Rendón *et al.*, 2007
ETEC	109/136 (80.1%)	25/43 (58.1%)[d]	Blackburn *et al.*, 2009
EPEC	30/30 (100%)	19/30 (63.3%)[e]	Saldaña *et al.*, 2009
atypical EPEC	29/29 (100%)	Not determined	Scaletsky *et al.*, 2010
atypical EPEC	61/71 (85.9%)	26/71 (36.6%)[f]	Hernandes *et al.*, 2011
EAEC	125 /130 (96.2%)	79/125 (63.2%)[f]	Avelino *et al.*, 2010
UPEC	28 / 28 (100%)	Not determined	Narciso *et al.*, 2011
Total	576/627 (91.9%)	274/465 (59.1%)	

[a] EAEC, enteroaggregative *E. coli*; EPEC, enteropathogenic *E. coli*; ETEC, enterotoxigenic *E. coli*; IPEC, intestinal pathogenic *E. coli*; NFEC, normal flora *E. coli*; NMEC, neonatal meningitis *E. coli*; UPEC, uropathogenic *E. coli*.

[b] Detected in cells cultivated at 20°C in Luria-Bertani (LB) by indirect immunofluorescence microscopy (IFM) and serum agglutination.

[c] Detected in cells cultivated at 26°C in Dulbecco's Modified Eagle Medium (DMEM) by flow cytometry.

[d] Detected in cells cultivated at 37°C under 5% CO_2 in pleuropneumonia-like organism (PPLO) medium by Western blotting.

[e] Detected in cells cultivated at 37°C in DMEM with cultured HT-29 colonic epithelial cells by flow cytometry and IFM.

[f] Detected in cells cultivated at 37°C in DMEM with cultured HEp-2 larynx epithelial cells by IFM.

Several studies support the idea that the Mat/ECP fimbria is an important colonization factor of *E. coli*. Various *E. coli* isolates from diverse origins are able to produce the fimbriae when adhering to cultured human epithelial cells (Avelino *et al.*, 2010; Blackburn *et al.*, 2009; Hernandes *et al.*, 2011; Rendón *et al.*, 2007; Saldaña *et al.*, 2009), and functional analysis of isogenic *matB* (*ecpA*) mutants have confirmed a role for Mat/ECP in host-cell adherence (Avelino *et al.*, 2010; Lasaro *et al.*, 2009; Rendón *et al.*, 2007; Saldaña *et al.*, 2009).

Recently, Garnett *et al.* (2012) showed that EcpD/MatE is preferentially located at the tip of the fimbrial stalk and performs the role of a tip-associated adhesin. Further, available data indicate that the Mat/ECP fimbria is expressed *in vivo*. Rendón *et al.* (2007) were able to detect circulating antibodies against Mat/ECP in different pools of sera, collected from bovines, healthy humans, and patients suffering hemolytic uremic syndrome caused by EHEC. Lasaro and co-workers (2009) demonstrated that the fimbria is an essential colonization determinant for the probiotic *E. coli* strain Nissle 1917 in the intestine of infant mouse. However, the tissue receptor for Mat fimbriae remains to be identified.

The regulation of Mat fimbriae is modulated by growth temperature and the composition of the culture medium, and the expression seems to differ among clonal groups of *E. coli*. While NMEC strains produced the fimbriae in LB broth at 20°C (Pouttu *et al.*, 2001), growth in the eukaryotic-cell culture medium DMEM in the presence of 5% CO_2 induced the Mat expression in IPEC both at 26°C and 37°C (Table 2). In contrast, cultivation in LB did not induce the production of Mat fimbriae in EHEC at 26°C and 37°C or EPEC at 37°C (Rendón *et al.*, 2007; Saldaña *et al.*, 2009). Numerous genome-wide mRNA analyses (Table 3) and studies of protein-DNA interactions (Cho *et al.*, 2008; Grainger *et al.*, 2005; Oshima *et al.*, 2006; Shimada *et al.*, 2011) indicate that *mat* expression is controlled by several global regulators, including H-NS, CRP-cAMP, and leucine-responsive protein Lrp. In addition, immunofluorescence microscopy analysis of NMEC by Pouttu *et al.* (2001) suggests that expression of Mat fimbriae is regulated by phase variation. However, detailed information on the influence of regulatory systems on *mat* expression is not yet available.

Table 3. *mat* genes in transcriptome analyses.

Study set-up	Strain background	Differentially regulated *mat* gene(s)	Fold up-regulation	Reference
hns mutation (Tn5seq1)	FB8 (K-12)	*matBCD*	1.7-8.3	Hommais *et al.*, 2001
deletion of *hns*	536 (UPEC)	*matABCDEF*	2.3-13.7	Müller *et al.*, 2006
RpoE (σE) overexpression	MG1655Z1 (K-12)	*matA*	3.8	Bury-Moné *et al.*, 2009
Acute acid stress (pH 3)	86-24 (EHEC O157□:□H7)	*matBCDEF*	4.1-28.8	House *et al.*, 2009
deletion of *gadE*	Sakai (EHEC O157□:□H7	*matB*	1.3	Kailasan Vanaja *et al.*, 2009
20 mM butyrate vs 20 mM L-leucine	Sakai (EHEC O157□:□H7)	*matABCDEF*	7.8-22.0	Tobe *et al.*, 2011

2 Aims of the study

At the time this thesis was initiated, the Mat fimbria of *E. coli* had been identified, but its functions and regulatory mechanisms were not characterized. The first publication characterizing the Mat fimbria (Pouttu *et al.*, 2001) described a particularly exciting observation: the major subunit gene was very common among *E. coli* isolates of diverse origin, in fact more common than the "common" type 1 fimbria, but the surface expression of the Mat fimbria under routine laboratory conditions was restricted to a highly virulent pathogroup of *E. coli*, i.e. the NMEC O18ac:K1:H7 isolates. Resolution of the molecular basis and evolutionary aspects of this differential expression and its possible adaptive advantages for NMEC formed the main motivation for my PhD study. During this thesis work, several studies (e.g. Lasaro *et al.*, 2009; Rendón *et al.*, 2007; see Table 2) expanded the spectrum of Mat-expressing strains beyond the NMEC isolates. The studies provided information suggesting that Mat regulation may differ in *E. coli* pathovars and that the function(s) of Mat fimbria is(are) common, and hence probably important for *E. coli* as a species.

The specific aims of the present study were:

1) to reveal a biological function(s) for the Mat fimbria;
2) to characterize the *mat* operon genes;
3) to detect environmental cues that promote the expression of Mat fimbriae;
4) to identify components of the regulatory circuit(s) controlling the expression of *mat* operon; and
5) to resolve the molecular mechanisms behind the differential *mat* expression in NMEC and other *E. coli* pathogroups.

3 Materials and methods

3.1 Bacterial strains and plasmids

The strains and plasmids used in this study are listed in Table 4 and Table 5, respectively.

Table 4. Bacterial strains used in this study.

Bacterial strain	Description	Article	Source or reference
Escherichia coli			
IHE 3034	clinical isolate, NMEC, O18:K1:H7, (also known as RK261)	II	Achtman *et al.*, 1983; Korhonen *et al.*, 1985
IHE 3034-113	IHE 3034 *rne131*	-	This study
IHE 3034-Rif	Rifampin-resistant IHE 3034	I-IV	Pouttu *et al.*, 2001
IHE 3034-89	IHE 3034-Rif Δ*matA* (+1 to +591 bp from the GTG of *matA*)	-	This study
IHE 3034-90	IHE 3034-Rif *matB*::*cat*	I	Pouttu *et al.*, 2001
IHE 3034-91	IHE 3034-Rif Δ*matB* (+61 to +528 bp from the ATG of *matB*)		Lehti *et al.*, 2010
IHE 3034-92	IHE 3034-Rif Δ*matC* (+61 to +609 bp from the ATG of *matC*)	I	This study
IHE 3034-93	IHE 3034-Rif Δ*matD* (+61 to +2466 bp from the ATG of *matD*)	I	This study
IHE 3034-94	IHE 3034-Rif Δ*matE* (+61 to +1584 bp from the ATG of *matE*)	I	This study
IHE 3034-95	IHE 3034-Rif Δ*matF* (+151 to +696 bp from the ATG of *matF*)	I	This study
IHE 3034-96	IHE 3034-Rif Δ*matBCDEF* (+61 bp from the ATG of *matB* to +753 bp from the ATG of *matF*)	I, III, IV	This study
IHE 3034-97	IHE 3034-Rif Δ*matBCDEF* complemented with *matBCDEF*	I	This study
IHE 3034-101	IHE 3034-Rif *matA*(G496C), resulting in the substitution MatA A166P	II	This study
IHE 3034-102	IHE 3034-Rif *matA*(A536C), resulting in the substitution MatA H179P	II, III, IV	This study
IHE 3034-103	IHE 3034-102 complemented with *matA*	II	This study
IHE 3034-104	IHE 3034-Rif *rrnB* $T1_4$-*lacZ*	II	This study
IHE 3034-105	IHE 3034-Rif *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +81 from the GTG of *matA* (*pA*)]	II	This study
IHE 3034-107	IHE 3034-Rif *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +745 from the GTG of *matA*(A536C) (*pAB*)]	II	This study
IHE 3034-108	IHE 3034-102 *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +81 from the GTG of *matA*(A536C) (*pA*)]	II	This study
IHE 3034-110	IHE 3034-102 *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +745 from the GTG of *matA*(A536C) (*pAB*)]	II	This study
IHE 3034-120	IHE 3034-Rif *rrnB* $T1_4$-*prcsB-lacZ* (-803 to +20 from the ATG of *rcsB*)	III	This study
IHE 3034-121	IHE 3034-Rif *rrnB* $T1_4$-*prcsD-lacZ* (-259 to +16 from the ATG of *rcsD*)	III	This study
IHE 3034-122	IHE 3034-102 *rrnB* $T1_4$-*prcsB-lacZ* (-803 to +20 from the ATG of *rcsB*)	III	This study
IHE 3034-123	IHE 3034-102 *rrnB* $T1_4$-*prcsD-lacZ* (-259 to +16 from the ATG of *rcsD*)	III	This study

IHE 3034-130	IHE 3034-Rif *rrnB* $T1_4$-*pflhD-lacZ* (-942 to +22 from the ATG of *flhD*)	IV	This study
IHE 3034-131	IHE 3034-Rif rrnB $T1_4$-*pfliC-lacZ* (-712 to +38 from the ATG of *fliC*)	IV	This study
IHE 3034-132	IHE 3034-102 *rrnB* $T1_4$-*pflhD-lacZ* (-942 to +22 from the ATG of *flhD*)	IV	This study
IHE 3034-133	IHE 3034-102 *rrnB* $T1_4$-*pfliC-lacZ* (-712 to +38 from the ATG of *fliC*)	IV	This study
IHE 3034-S1	IHE 3034-Rif *S1* promoter region (-167 to +10 from the GTG of *matA*) swapped from MG1655	II	This study
IHE 3034-S2	IHE 3034-Rif *S2* promoter region (-343 to -147 from the GTG of *matA*) swapped from MG1655	II	This study
IHE 3034-S3	IHE 3034-Rif *S3* promoter region (-443 to -323 from the GTG of *matA*) swapped from MG1655	II	This study
IHE 3034-S4	IHE 3034-Rif *S4* promoter region (-608 to -421 from the GTG of *matA*) swapped from MG1655	II	This study
IHE 3034-S12	IHE 3034-Rif *S12* promoter region (-343 to +10 from the GTG of *matA*) swapped from MG1655	II	This study
IHE 3034-S34	IHE 3034-Rif *S34* promoter region (-608 to -323 from the GTG of *matA*) swapped from MG1655	II	This study
IHE 3034-150	IHE 3034-Rif Δ*crp* (+6 to +630 bp from the ATG of *crp*)	-	This study
IHE 3034-Sm	Streptomycin-resistant IHE 3034, O18:K1:H7	I, II, III	Pouttu *et al.*, 1999
IHE 3034-2	IHE 3034-Sm *fimA*::*cat*	I	Pouttu *et al.*, 1999
IHE 3034-8	IHE 3034-Sm *sfaA*::Gm	I	Pouttu *et al.*, 1999
IHE 3034-80	IHE 3034-Sm *fliC*::*cat*	I	Pouttu *et al.*, 1999
IHE 3034-59	IHE 3034-Sm *fimA*::*cat sfaA*::Gm		This study
IHE 3034-79	IHE 3034-Sm *fimA*::*cat sfaA*::Gm *fliC*::*cat*	I, IV	Pouttu *et al.*, 2001
IHE 3034-99	IHE 3034-Sm *fimA*::*cat sfaA*::Gm *fliC*::*cat* Δ*matB*	I	This study
IHE 3034-200	IHE 3034-Sm *hns*::mini-Tn5*cat*	II	This study
IHE 3034-201	IHE 3034-Sm *hns*::mini-Tn5*cat* *matA*(G496C), resulting in the substitution MatA A166P	II	This study
IHE 3034-202	IHE 3034-Sm *hns*::mini-Tn5*cat* *matA*(A536C), resulting in the substitution MatA H179P	II	This study
IHE 3034-203	IHE 3034-Sm *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +81 from the GTG of *matA* (*pA*)]	II, III	This study
IHE 3034-204	IHE 3034-Sm *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +745 from the GTG of *matA*(A536C) (*pAB*)]	II, III	This study
IHE 3034-205	IHE 3034-200 *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +81 from the GTG of *matA* (*pA*)]	II	This study
IHE 3034-206	IHE 3034-200 *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +745 from the GTG of *matA*(A536C) (*pAB*)]	II	This study
IHE 3034-207	IHE 3034-202 *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +81 from the GTG of *matA* (*pA*)]	II	This study
IHE 3034-208	IHE 3034-202 *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +745 from the GTG of *matA*(A536C) (*pAB*)]	II	This study

IHE 3034-213	IHE 3034-Sm *rcsB*::mini-Tn*5cat rrnB* $T1_4$-*pmatA-lacZ* [-608 to +81 from the GTG of *matA* (*pA*)]	III	This study
IHE 3034-214	IHE 3034-Sm *rcsB*::mini-Tn*5cat rrnB* $T1_4$-*pmatA-lacZ* [-608 to +745 from the GTG of *matA*(A536C) (*pAB*)]	III	This study
IHE 3034-215	IHE 3034-Sm *rrnB* $T1_4$-*pmatA-lacZ* [-608 to +745 with $G^{-340}T^{-337}C^{-336}A^{-331}G^{-330}C^{-327}$ from the GTG of *matA*(A536C) ($pAB_{RcsAB\ box\ mut}$)]	III	This study
IHE 3034-216	IHE 3034-204 Δ*rcsC* (+61 to +2790 bp from the TTG of *rcsC*)	III	This study
IHE 3034-217	IHE 3034-204 Δ*rcsD* (+61 to +1834 bp from the ATG of *rcsD*]	III	This study
IHE 3034-218	IHE 3034-204 Δ*ackA* Δ*pta* (+61 to +3362 bp from the ATG of *ackA*)	III	This study
MG1655	K-12, OR:H48	I, II	Blattner *et al.*, 1997
MG1655-Rif	Rifampin-resistant MG1655	II, IV	This study
MG1655-102	MG1655-Rif *matA*(A536C), resulting in the substitution H179P	II, IV	This study
MG1655-S12	MG1655-Rif *S12* promoter region (-343 to +10 from the GTG of *matA*) swapped from IHE 3034	II	This study
MG1655-S123	MG1655-Rif *S123* promoter region (-443 to +10 from the GTG of *matA*) swapped from IHE 3034	II	This study
MG1655-S1234	MG1655-Rif *S1234* promoter region (-608 to +10 from the GTG of *matA*) swapped from IHE 3034	II	This study
AES29	MG1655 *hns*::*cat*	II	A. Sjöström and B.E. Uhlin, Umeå University, unpublished construction
HB101	K-12	-	American Type Culture Collection
IHE 3040	clinical isolate, NMEC, O18:K1:H7	II	Achtman *et al.*, 1983; Korhonen *et al.*, 1985
IHE 3072	clinical isolate, NMEC, O2:K1:H5	II	Achtman *et al.*, 1983; Korhonen *et al.*, 1985
BEN79	APEC, O18:K1:H7	II	Moulin-Schouleur *et al.*, 2007
BEN374	APEC, O18:K1:H7	II	Moulin-Schouleur *et al.*, 2007
536	clinical isolate, UPEC, O6:K15:H31	II	Berger *et al.*, 1982
CFT073	clinical isolate, UPEC, O6:K2:H1	II	Mobley *et al.*, 1990
Nissle 1917	probiotic strain, DSM 6601, O6:K5:H1	II	Grozdanov *et al.*, 2004
83972	clinical isolate, ABU, OR:K5:H–	II	Lindberg *et al.*, 1975
E2348/69	clinical isolate, EPEC, O127:H6	II	Levine *et al.*, 1978
LF82	clinical isolate, AIEC, O83:H1	II	Darfeuille-Michaud *et al.*, 1998
ABU 38	clinical isolate, ABU	I, II	Lindberg *et al.*, 1975
285	clinical isolate, NMEC, O78	II	Ron *et al.*, 1991
789	APEC, O78	II	Babai *et al.*, 1997
BL21(DE3)	F^- *ompT dcm lon hsdS*$_B$ (r_B^- m_B^-) *gal* λDE3	II, IV	Studier & Moffatt, 1986
BL21(AI) *slyD*	F^- *ompT dcm hsdS*$_B$ ($r_B^-m_B^-$) *gal araB*::T7RNAP-*tetA* Δ*slyD*::*cat*	III	Daniel *et al.*, 2006
S17-1 λpir	*recA thi pro hsdR* RP4-2-Tc::Mu Km::Tn7 Sm^R Tp^R λ*pir*	I-IV	Simon *et al.*, 1983

Sm10 λpir	*thi*, *thr*, *leu*, *tonA*, *lacY*, *supE*, *recA*::RP4-2-Tc::Mu Km^R λpir	II, III	Simon *et al.*, 1983
Bifidobacterium longum			
ATCC 15707	stool isolate from adult	II	American Type Culture Collection

Table 5. Plasmids used in this study.

Plasmid	Description	Article	Source or reference
pSE380	Expression vector, *ptrc*, *lacI, lacO*, Ap^R	I-IV	Invitrogen
pMAT19	IHE 3034 *matA* under *ptrc* in pSE380	II, III, IV	This study
pMAT10	IHE 3034 *matB* under *ptrc* in pSE380	I	Pouttu *et al.*, 2001
pMAT12	IHE 3034 *matC* under *ptrc* in pSE380	I	This study
pMAT13	IHE 3034 *matD* under *ptrc* in pSE380	I	This study
pMAT14	IHE 3034 *matE* under *ptrc* in pSE380	I	This study
pMAT15	IHE 3034 *matF* under *ptrc* in pSE380	I	This study
pMAT6	IHE 3034 *matBCDEF* under *ptrc* in pSE380	I, III, IV	Pouttu *et al.*, 2001
pMAT20	MG1655 *matBCDEF* under *ptrc* in pSE380	I	This study
pMAT21	MG1655 *matA* under *ptrc* in pSE380	II	This study
pRCSB1	IHE 3034 *rcsB* under *ptrc* in pSE380	III	This study
pRS551	Transcriptional *lacZ* fusion vector, Ap^R, Km^R		Simons *et al.*, 1987
pRS551-Δtrpt	pRS551 Δ*trpt* (+17 to +69 bp from the *Eco*RI site)	II	This study
pRS551-pA	IHE3034 *pmatA* in pRS551 [-608 to +81 from the GTG of *matA* (*pA*)]	II	This study
pRS551-pB	IHE3034 *pmatB* in pRS551 [-608 to +81 from the ATG of *matB* (*pB*)]	II	This study
pRS551-pC	IHE3034 *pmatC* in pRS551 [-615 to +81 from the ATG of *matC* (*pC*)]	II	This study
pRS551-pD	IHE3034 *pmatD* in pRS551 [-608 to +81 from the ATG of *matD* (*pD*)]	II	This study
pRS551-pE	IHE3034 *pmatE* in pRS551 [-612 to +83 from the ATG of *matE* (*pE*)]	II	This study
pRS551-pF	IHE3034 *pmatF* in pRS551 [-608 to +81 from the ATG of *matF* (*pF*)]	II	This study
pRS551-pAΔtrpt	IHE3034 *pmatA* in pRS551-Δtrpt [-608 to +81 from the GTG of matA (*pA*-Δ*trpt*)]	II	This study
pRS551-pP1	IHE3034 *pmatA* in pRS551 [-121 to -56 from the GTG of *matA* (*pP1*)]	II	This study
pRS551-pP2	IHE3034 *pmatA* in pRS551 [-343 to -122 from the GTG of *matA* (*pP2*)]	II	This study
pRS551-pP3	IHE3034 *pmatA* in pRS551 [-608 to -344 from the GTG of *matA* (*pP3*)]	II	This study
pRS551-pP12	IHE3034 *pmatA* in pRS551 [-343 to -56 from the GTG of *matA* (*pP12*)]	II	This study
pRS551-pP23	IHE3034 *pmatA* in pRS551 [-608 to -122 from the GTG of *matA* (*pP23*)]	II	This study
pRS551-pP123	IHE3034 *pmatA* in pRS551 [-608 to -56 from the GTG of *matA* (*pP123*)]	II	This study
pRS551-pA2	MG1655 *pmatA* in pRS551 [-608 to +81 from the GTG of matA (pA_{MG1655})]	II	This study
pUC19	cloning vector, Ap^R	II	Yanisch-Perron *et al.*, 1985
pUC19-placZ	IHE3034 *lacI*-pRS551 *rrnB* $T1_4$-partial	II, III, IV	This study

	lacZ in pUC19		
pUC19-pAB	IHE3034 *pmatA* [-608 to +745 from the GTG of *matA*(A536C)] in pUC19-placZ	III	This study
pACYC184	cloning vector, Tc^R, Cm^R	II, IV	New England Biolabs
pTAL1	IHE 3034 *matA* under *ptet* in pACYC184	II, IV	This study
pTAL2	MG1655 *matA* under *ptet* in pACYC184	II, IV	This study
pBR322	cloning vector, Tc^R, Ap^R	II	Bolivar *et al.*, 1977
pHNS1	IHE 3034 *hns* with 660 bp upstream and 143 bp downstream region in pBR322	II	This study
pHNS2	MG1655 *hns* with 661 bp upstream and 143 bp downstream region in pBR322	II	This study
pCRP1	IHE 3034 *crp* with 299 bp upstream region in pBR322	-	This study
pMAL-c2x	Maltose-binding protein (MBP) fusion vector, *malE*, *lacI*, *ptac,* Ap^R	II, III, IV	New England Biolabs
pBAU1	IHE 3034 *matA* in pMAL-c2x	II, III, IV	This study
pBAU1-H2	IHE 3034 *matA*(A536C) in pMAL-c2x	II	This study
pMAL-RCSA1	IHE 3034 *rcsA* in pMAL-c2x	III	This study
pQE30	N-terminal 6xHistag fusion vector, *pT5, lacO*, Ap^R	II, III	Qiagen
pQE30-HNS1	IHE 3034 *hns* in pQE30	II	This study
pQE30-RCSB1	IHE 3034 *rcsB* in pQE30	III	This study
pQE30-RCSB2	IHE 3034 *rcsB*(T168G) in pQE30, resulting in $RcsB_{D56E}$	III	This study
pQE30-CRP1	IHE 3034 *crp* in pQE30	-	This study
pQE30-MATA1	IHE 3034 *matA* in pQE30	-	R. Pouttu, University of Helsinki, unpublished construction
pQE60	C-terminal 6xHistag fusion vector, *pT5, lacO*, Ap^R	-	Qiagen
pQE60-MATA1	IHE 3034 *matA* in pQE60	-	This study
pREP4	Repressor plasmid, *lacI*, Km^R	II, III	Qiagen
pGEX-6p-1	Glutathione S-transferase (GST) fusion vector, *lacI*, *ptac*, Ap^R	-	Amersham Biosciences
pGEX-MATA1	IHE 3034 *matA* in pGEX-6p-1	-	This study
pGP704	Suicide vector, pir-dependent *ori* R6K, *mob* RP4, Ap^R	I	Miller & Mekalanos, 1988
pCVD442	Suicide vector, *pir*-dependent *ori* R6K, *mob* RP4, *sacB*, Ap^R	I-IV	Donnenberg & Kaper, 1991
pKO3	Suicide vector, temperature-sensitive *oripSC101* replication origin, *sacB*, Cm^R	-	Link *et al.*, 1997
pKD4	*kan* cassette template vector for RED system, Ap^R, Km^R	-	Datsenko & Wanner, 2000
pKD46	λ Red recombinase expression plasmid for RED system, temperature-sensitive *oriR101 repA101* replication origin*, γ-β-*exo under *paraBAD*, Ap^R	-	Datsenko & Wanner, 2000
pCP20	FLP recombinase expression plasmid for RED system, temperature-sensitive *oripSC101*, cI857λ, *flp* under λ p_R, Ap^R, Cm^R	-	Cherepanov & Wackernagel, 1995
pUTmini-Tn5 Cm	Transposon vector, *pir*-dependent *ori* R6K, *mob* RP4, Ap^R, Cm^R	II, III	Biomedal
pGEM®-T Easy	Cloning vector, Ap^R	II	Promega

3.2. Methods used in this thesis

The experimental procedures used in this study are described in detail in the original publications, and compiled in Table 6; methods not used in the publications are described in the Chapter 3.3.

Table 6. Summary of methods used in this study.

Method	Article
Standard DNA techniques	I-IV
Site-directed mutagenesis	I-IV
Transposon mutagenesis	II, III
DNA sequence analysis	II, III
RNA isolation	II, III, IV
Northern blotting	II, III, IV
mRNA half-life determination	II
RT-PCR	II
Real-time RT-PCR	II
Primer extension	II
5' RACE	II
Whole-cell ELISA	I, II
Indirect immunofluorescence	I
Western blotting	I, II, IV
Colony immunoblotting	II
Screening with immunomagnetic particles	II
Serum agglutination	I, II, III
Protein expression and purification	II, III, IV
Electrophoretic mobility shift assay	II, III, IV
N-terminal sequencing	II
β-galactosidase assay	II, III, IV
Biofilm formation assay	I, III
Initial attachment assay	I
Motility assay	IV

3.3 Methods other than those described in the papers I - IV

3.3.1 Hemagglutination and yeast cell agglutination assays

Agglutination of *Saccharomyces cerevisiae* yeast cells or the human group O erythrocytes (Reagent red cells for antibody identification - DiaMed system, Alba Bioscience, UK; contains 11 different untreated and 11 papain-treated human group O red cells) by bacteria were performed as previously described (Korhonen & Finne, 1985) by mixing equal volumes of a suspension of PBS-washed bacteria and eukaryotic cells on cold glass slides placed on ice. The mixture was gently rotated for 5 min, after which bacterial agglutination was detected visually, and confirmed by phase-contrast microscopy using an Olympus CKX41 microscope.

3.3.2 Interactions with human epithelial cell lines

The human cell lines used in the present study are Henle 407 (embryonic intestine, ATCC CCL-6), T24 (human bladder carcinoma), VK2/E6E7 (vaginal, ATCC CRL-2616), Ect1/E6E7 (ectocervical, ATCC CRL-2614), and End1/E6E7 (endocervical, ATCC CRL-2615). Henle 407 were routinely cultured in Basal medium Eagle supplemented with 10% fetal bovine serum and 2 mM L-Glutamine; T24 cells were cultured in McCoy's 5A medium supplemented with 10% fetal bovine; and VK2/E6E7, Ect1/E6E7 and End1/E6E7 were cultured in keratinocyte serum-free medium supplemented with 0.1 ng ml^{-1} human recombinant EGF, 0.05 mg ml^{-1} bovine pituitary extract, and additional calcium chloride (final concentration 0.4 nM). Cells were maintained at 37°C in 5% CO_2.

For adhesion assays, the eukaryotic cells were grown on eight-well chamber slides (Lab-Tek® II, Nalge Nunc International) or diagnostic eight-well glass slides for 1-2 days to approximately 70% confluence. The culture medium was removed and cells were fixed with cold 3.5% paraformaldehyde in PBS for 10 min or alternatively left untreated. After washing three times with PBS, bacterial suspension (40 µl, 5×10^8, 10^9 or 5×10^9 ml^{-1} in PBS, in respective eukaryotic-cell culture medium, or in DMEM) were added on the eukaryotic cells and incubated for 1-3 hour at 37°C, in 5% CO_2. The slides were washed five times with PBS, fixed with methanol for 10 min, dried, and stained with 0.1% Giemsa stain. The adherent bacteria were examined in an Olympus BX50 or Olympus CKX41 microscope.

3.3.3 Immunoelectron microscopy

For immunoelectron microscopy, bacteria were grown overnight at 37°C in LB broth buffered with 100 mM MES (2-(N-Morpholino)ethanesulfonic acid), pH 5.5. Isopropyl-β-D-thiogalactopyranoside (IPTG) and ampicillin, when necessary, were added at concentrtions of 5 µM and 100 µg ml^{-1}, respectively. The bacteria were collected by centrifugation (2000 rpm 10 min) and suspended in LB broth. The bacteria were prepared for immunoelectron microscopy as described by Pouttu *et al.* (2001). Briefly, the bacteria were transferred to copper grids coated with Pioloform and carbon by incubating for 10 min. The immobilized bacteria were left to react for 45 min with anti-Mat antibodies diluted to 1:500 in PBS containing 0.1% BSA, and then for 45 min with 10-nm-particle-diameter gold-labeled protein A (Amersham Life Science) diluted to 1:65 in PBS containing 0.1% BSA. For contrasting the grids, the bacteria were negatively stained by floating on drops of 1% phosphotungstic acid in water, pH 6.5 for 1.5 min. After removal of excess fluid, microscopy was done with a JEOL JEM-1200EX transmission electron microscope operated at 60 kV, at Electron Microscopy Unit, Institute of Biotechnology, University of Helsinki.

3.3.4 Prediction of signal peptides

The presence of signal peptide cleavage sites in amino acid sequences were predicted by SignalP v3.0 based on hidden Markov models (Bendtsen *et al.*, 2004).

3.3.5 Frameshift mutation of *rne* gene

In order to analyze the effects of RNase E inactivation on temperature-dependent *mat* expression, the *rne131* mutant allele encoding a stable, truncated form of RNase E (Kido *et al.*, 1996) was chosen instead of widely used temperature-sensitive *rne* alleles (*rne-1* and *rne-3071*) which utilization requires compulsory temperature upshift that affects the endoribonucleolytic activity of the encoded mutant proteins (McDowall *et al.*, 1993). To create the *rne131* mutation into the IHE 3034 chromosome, a deletion of two nucleotides (2295GT) resulting in a +1 frameshift was introduced into the *rne* gene. Briefly, the 5' and 3' flanking sequences of the deletion site were amplified from IHE 3034 chromosomal DNA with the upstream primer 252F 5'-GTAGCGGCCGCAGATCGATAAAGAAGAGC-3' and the mutagenic primer 247R 5'-TGGTTTCTTCCACCGCTGAACAGCGC-3'; and the mutagenic primer 246F 5'-GTAGTCGACTTACTCAACAGGTTGCGG-3' and the downstream primer 253R 5'-TTCAGCGGTGGAAGAAACCAAACCGACC-3'. The PCR products were fused with each other by recombinant PCR and then cloned into the *Not*I-*Sal*I site of pKO3. A previously described gene replacement method which utilizes the pKO3 plasmid with a temperature-sensitive origin of replication (Link *et al.*, 1997) was applied, and the created mutation was verified by DNA sequencing using primer 251F 5'-CCAACCTTAAGCTACATGC-3'.

3.3.6 Deletion of *matA* gene

In-frame deletion of *matA* on the chromosome of the strain IHE 3034-Rif was created by site-specific mutagenesis using recombinant PCR and the *pir*-dependent suicide vector pCVD442 essentially as described previously (Mobley *et al.*, 1993). Briefly, the 5' and 3' flanking sequences of the deletion site were amplified with the upstream primer 003F 5'-GCAGAGCTCATATCCATCCTGAGTATCG-3' and the mutagenic primer 006R 5'-CTTGCCTGGAGTTACTTTCCAAACCTGTAATTT-3'; and the mutagenic primer 005F 5'- GTTTGGAAAGTAACTCCAGGCAAGTTAG-3' and the downstream primer 004R 5'-GCAGCATGCGTCGTAAAGGCATTGTCG-5'. The overlapping PCR products were purified and mixed, and a fusion product was digested with *Sac*I and *Sph*I, and cloned into pCVD442. The resulting suicide plasmid was mobilized into IHE 3034-Rif via conjugation from S17-1 λpir. After sucrose selection, the allelic exchange was screened by colony-PCR, and the correct deletion was verified by DNA sequencing using primer 014F 5'-CATATTGACACTCATCAGG-3'.

3.3.7 Deletion and complementation of *crp* gene

The deletion of *crp* in the chromosome of IHE 3034-Rif was generated by using the λ Red-based technology and linear DNA for recombination, as described previously by Datsenko and Wanner (2000). The kanamycin *kan* gene flanked by FLP recognition target (FRT) sites, from the plasmid pKD4, was amplified by PCR using primers 175REDF 5'-CTCTGGAGAAAGCTTATAACAGAGGATAACCGCGCATGGTGTAGGCTGGAGCTG CTTC-3' and 176REDR 5'-GGCGCGCTACCAGGTAACGCGCCACTCTGACGGGATT ACATATGAATATCCTCCTTAG-3' (58 nt) containing 38 nucleotides of homologous sequence to the region 5' and 3' of the *crp* gene. After electroporation of IHE 3034-Rif

(transformed with the pKD46 helper plasmid) by using 250 ng of *Dpn*I digested and purified PCR fragment, and exchange of the wild-type *crp* allele with the FRT-flanked *kan* insertion, the resistance gene was eliminated with the help of the FLP recombinase encoded on plasmid pCP20. The generated deletion was verified by DNA sequencing using primer 177F 5'- CACAAAGCGAAAGCTATGC-3'.

Plasmid pCRP1 was constructed by amplication of the *crp* coding region together with 299-bp upstream region from the strain IHE 3034 by using the oligonucleotide primers 179F 5'-CGCGGATCCTTGCTACTCCACTGCGTC-3' and 180R 5'-CGGCGAATTCAAATGGCG CGCTACCAGG-3', and cloning the resulting PCR product into the *Bam*HI-*Eco*RI restriction sites of vector pBR322.

3.3.8 Purification of His_6-MatA and GST-MatA proteins

Plasmid pQE60-MATA1 was constructed by amplication of the *matA* coding region from the strain IHE 3034 by using primers 117F 5'-GCTCCATGGGAACATGGCAAAATGATT ACAG-3' and 123R 5'-CGCAGATCTCTGAACCAACTTATATAATTT-3', and cloning the resulting PCR product into the *Nco*I-*Bgl*II restriction sites of vector pQE60.

To express glutathione S-transferase (GST) -MatA fusion protein, a PCR-generated fragment encoding MatA from IHE 3034 was amplified by 165F 5'-GCCGGATCCGTGACATGGCA AAATGATTAC-3' and 008R 5'-GCCGTCGACTTACTGAACCAACTTATATAAT-3' primers, and cloned into *Bam*HI-*Sal*I digested pGEX-6p-1, to generate pGEX-MATA1.

N- and C-terminally His_6-tagged MatA proteins were purified from BL21(AI) *slyD* (pREP4) cells harbouring expression plasmid pQE30-MATA1 or pQE60-MATA1, and GST-MatA from BL21(DE3) (pGEX-MATA1) cells. The cells were cultivated in LB medium at various temperatures (20, 30 or 37°C) to mid-logarithmic phase (corresponding to an OD600 of around 0.5) at which point the culture temperature was reduced to 15°C or kept unchanged, and IPTG was added to a final concentration of 0.1 or 1 mM. After 4 h incubation, the cells were harvested by centrifugation and frozen at -70°C. The His-tagged recombinant proteins were purified under native or denaturing conditions using Ni-nitrilotriacetic acid resin according to the QIA *express* System (QIAGEN). GST-MatA was purified using Gluthatione Sepharose™ 4B according to manufacturer's instructions (GE Healthcare).

3.3.9 Purification of His_6-CRP protein and electrophoretic mobility shift assay

To express His_6-CRP protein protein, a 0.6-kb BamHI/HindIII DNA fragment containing the *crp* coding sequence from IHE 3034 was amplified by 181F 5'-CGCGGATCCATGGTGCT TGGCAAACCG-3' and 182R 5'-CGCGAAGCTTTTAACGAGTGCCGTAAACG-3' primers and inserted between the *Bam*HI and *Hind*III restriction sites of the pQE30 vector, yielding the plasmid pQE30-CRP1.

BL21(DE3) (pQE30-CRP1)(pREP4) cells were grown in 600 ml of LB media with appropriate antibiotics to OD_{600} of 0.7 at 30°C, and expression of the fusion protein was then induced for 4 h after the addition of 1 mM IPTG. Cells were harvested by **c**entrifugation and frozen at -70°C. The His-tagged CRP proteins were purified under native conditions using

Ni-nitrilotriacetic acid resin according to the QIA express System (QIAGEN). Protein-containing elution fractions, as judged by A_{280} absorption, were loaded into Spectra/Pro® dialysis membrane tube (12-14 kDa cut-off; Spectrum Laboratories) and step-wise dialysed into 30 mM Tris-HCl at pH 7.5–150 mM NaCl–15% glycerol. The protein concentrations were estimated from SDS-PAGE gels comparing intensities of stained polypeptides with those of BSA standards of known concentration using Tina (v2.0) image analysis program (Raytest Isotopenmessgeräte Gmbh).

The *pmatA, pmatB2* and *placZ* DNA fragments used for the EMSA assay were obtained by PCR amplification using genomic DNA from IHE 3034 as a template. The primers for *pmatA* are listed in the Table S4 of Paper II. The *pmatB2* fragment was amplified by using primers MATB-F 5'-ATGAAAAAAAAGGTTCTGGCA-3 and 114R 5'-GTAGCCGTTGAGTCACC-3', and *placZ* using 191F 5'-TTAATGCAGCTGGCACGAC-3' and 190R 5'-GGCCAGTGAATCCGTAATC-3'. An equimolar mixture of the fragments (0.28 pmol of each) were incubated with increasing concentrations of His_6–CRP in 10-µl reaction mixtures 10 mM Tris-HCl at pH 7.5–1 mM EDTA–50 mM NaCl–5 mM $MgCl_2$–1 mM dithiothreitol–0.05% Nonidet P-40–5% glycerol-0.1 mM cAMP. Following an incubation for 15 min at 20°C, the reactions were electrophoresed in 6% PAGE in 0.5×TBE at 20°C. After electrophoresis, the DNA fragments were visualized by staining with ethidium bromide and exposing to UV light.

3.4 Accession numbers of the nucleotide sequences

Table 7. Genbank accession numbers of the *E. coli* nucleotide sequences described in this study.

Sequence designation	Accession number	Article
mat operon region of IHE 3034	HM102365	I
mat upstream regulatory region of IHE 3040	JN377374	II
mat upstream regulatory region of IHE 3072	JN377375	II
mat upstream regulatory region of BEN79	JN377376	II
mat upstream regulatory region of BEN374	JN377377	II
mat upstream regulatory region of Nissle 1917	JN377378	II
mat upstream regulatory region of 285	JN377379	II
mat upstream regulatory region of 789	JN377380	II
matA coding region of BEN374	JN377381	-

4 Results and discussion

In this thesis work, the O18ac:K1:H7 meningitis isolate *E. coli* IHE 3034 was selected as the model organism to study functions of Mat fimbria and regulatory networks controlling the expression of the *mat* operon. The strain was originally isolated in Finland in 1976 from a case of fatal late onset neonatal meningitis (Korhonen *et al.*, 1985; Selander *et al.*, 1986), and has been thereafter further characterized in various studies in Helsinki (e.g. Pouttu *et al.*, 1999 and 2001) as well as by research groups elsewhere (e.g. Bauchart *et al.*, 2010; Dobrindt *et al.*, 2003; Moriel *et al.*, 2010; Nagy *et al.*, 2005; Nougayrède *et al.*, 2006; Sjöström *et al.*, 2009b; Wang & Kim, 2000). IHE 3034 belongs to one of the best characterized NMEC strains and to the virulence-associated phylogenetic group B2. To facilitate functional genetic studies on Mat fimbriae, we first sequenced the six-gene *mat* operon and its flanking sequences of the *E. coli* IHE 3034 chromosome (Genbank accession no. HM102365; Table 7). Moreover, the complete 5.11-Mb genome sequence of IHE 3034 consisting of 4,757 translatable open reading frames became available during the last part of my thesis work (Moriel *et al.*, 2010). To study the differential expression of *mat* genes, the laboratory K-12 strain MG1655, in which Mat expression had not been detected although the strain has a seemingly intact *mat* gene cluster (Pouttu *et al.*, 2001; and this thesis work), was analyzed for comparison. M1655 is an A-phylogroup strain and the first *E. coli* isolate whose genome (4.64 Mb) was sequenced (Blattner *et al.*, 1997). The IHE 3034 strain contains two extrachromosomal plasmids (Achtman *et al.*, 1983) that are absent in MG1655 strain, and carries 0.47 Mb more chromosomal DNA than MG1655. This DNA is primarily organized to various IHE 3034-specific genomic and pathogenicity islands that encode many candidate virulence factors, such as the S fimbria, the K1 capsule, type II and VI secretion systems, several siderophores, and the siderophore receptor IroN (Dobrindt *et al.*, 2003; Moriel *et al.*, 2010). On the other hand, 202 genes present in MG1655 were not detectable in the genome of IHE 3034 (Dobrindt *et al.*, 2003). However, the *mat* gene cluster in the two strains is located at the same chromosomal site ($304373\text{-}311212_{\text{IHE 3034}}$ and $303719\text{-}310560_{\text{MG1655}}$) in both strains and has a sequence identity of 97.8%.

4.1 The common colonization factor Mat fimbria promotes biofilm formation

4.1.1 Assessing potential Mat fimbriae-mediated adherence on eukaryotic cells (this thesis)

In search for potential target molecules for Mat fimbriae-mediated adhesion, I first focused on eukaryotic cells. Previous study had shown that *E. coli* expressing Mat fimbriae does not agglutinate human O group erythrocytes, neither untreated nor treated with endo-β-galactosidase, neuraminidase, trypsin, or pronase to expose cryptic carbohydrates on red blood cell surface (Pouttu *et al.*, 2001). Testing with mannan-rich yeast (*Saccharomyces cerevisiae*) cells and panels of untreated or papain-treated human O group erythrocytes carrying various blood type antigens (Rhesus, Kell, Duffy, Kidd, Lewis, MNS, P, Lu, Wr among others) gave similar result; Mat fimbriae-expressing K-12 recombinant bacteria and flagellum-, S-fimbria-, and type 1-fimbria-negative IHE 3034-79 did not agglutinate any of the cells tested (data not shown) suggesting that common carbohydrates of human and yeast cell surfaces are not recognized by the Mat fimbria. Thus, the approaches that were

successful in characterizing P, type 1, S, and Dr fimbriae binding did not apply to the Mat fimbria.

Next, it was set out to investigate whether Mat fimbria promotes adherence to cultured human epithelial cells. IHE 3034-Rif and isogenic Mat fimbria-deficient *matB*::*cat* derivative, as well as laboratory K-12 strain HB101 harbouring plasmid pMAT6, encoding Mat fimbriae from IHE 3034, or vector pSE380 were evaluated for their ability to adhere to immortalized human intestinal (Henle 407), bladder (T24), vaginal (VK2/E6E7), ectocervical (Ect1/E6E7), and endocervical (End1/E6E7) epithelial cell lines. Surface expression of Mat fimbriae did not induce significant adherence to any of the tested cell lines (data not shown). While this thesis work was in progress, two other research groups addressed the role of ECP fimbriae in host cell adhesion and showed that the fimbriae mediate adherence to human cervix (HeLa), colonic (HT-29), and larynx (HEp-2) epithelial cell (Avelino *et al.*, 2010; Lasaro *et al.*, 2009; Rendón *et al.*, 2007; Saldaña *et al.*, 2009). Differences in receptor specificity of the adhesive fimbria subunit (Mat versus ECP), bacterial strains (NMEC versus e.g. EPEC and EHEC), and cultured eukaryotic cells used, or experimental conditions may explain our divergent outcome.

4.1.2 Mat fimbriae encoded by *matABCDEF* gene cluster mediate temperature-dependent biofilm formation (I, II, III and this thesis)

It was next investigated whether Mat fimbria contributes to *E. coli* biofilm formation ability on inert surfaces. Using different target surfaces [hydrophilic glass and hydrophobic polyvinyl chloride (PVC), polystyrene, and polypropylene], different media (LB broth and M63 minimal medium supplemented with glucose and casamino acids) and temperature conditions (20 and 37°C), we found that the clinical isolate IHE 3034 formed a maximum biofilm on hydrophobic PVC surface when grown in M63 medium at 20°C (Fig. 2 of I). Contrary to IHE 3034, the Mat fimbriae-negative laboratory K-12 strain MG1655 was unable to establish biofilm under these conditions. No biofilm on PVC surface was detected with the strains when the growth temperature was elevated to 37°C.

I assessed the contribution of the entire *matBCDEF* region and each putative structural *mat* gene on temperature-dependent biofilm formation by IHE 3034. Earlier work by Pouttu and co-workers (2001) had shown that MatB is a major structural subunit of the Mat fimbria, and complementation assays of *matB*-deficient mutant also indicated involvement of *matA* and *matC* in the surface expression of the fimbria. The N-terminal signal sequence directs secretory proteins cross the cytoplasmic membrane to the periplasm via the general secretion pathway (Sec-pathway) (Natale *et al.*, 2008). In silico analysis using SignalP 3.0 (Bendtsen *et al.*, 2004) indicated the presence of a cleavable N-terminal signal sequence in *matB-F* gene products but not in the predicted MatA sequence (data not show). This suggests that MatC, MatD, MatE, and MatF are either a periplasmic or an outer membrane protein, and may therefore directly be dedicated to the assembly of the filamentous Mat structures, as MatB fimbrillin, whereas MatA may be a cytoplasmic protein. To assess the participation of *matB-F* genes both in Mat fimbria biogenesis and biofilm formation, I inactivated each of these genes and the entire *matBCDEF* locus by nonpolar in-frame deletion. Whole-cell ELISA using anti-Mat fimbria antibodies revealed a significant reduction in surface

expression of Mat fimbriae in the deletion derivatives compared with the wild-type parent strain, IHE 3034-Rif (in each case $p < 0.001$, unpaired *t* test; Fig. 1 of I). The defect in Mat fimbriation was rescued by plasmid-derived expression of the corresponding gene, or in the case of *matB-F* deletion strain by insertion of the wild-type alleles back into the native chromosomal location. Functional analysis of the knockout mutants showed a complete loss of biofilm formation on PVC without affecting planktonic growth (Fig. 2 of I). In line with our results, a recent biofilm mutant screen on the EHEC O157:H7 strain EDL933 identified *matE* (or *ecpD*) to be involved in biofilm formation on polystyrene surface (Puttamreddy *et al.*, 2010). Similarly, another study showed the critical role of *matB* (or *ecpA*) for the ability of EDL933 to form biofilm (Garnett *et al.*, 2012).

Time course analysis of surface attachment demonstrated that Mat-defective cells were capable of reversible binding on PVC but unable to form a dispersed monolayer (Fig. 4 of I). The other known *in vitro* expressed cell-associated appendages of IHE 3034, type 1 fimbria, S fimbria, and flagellum (Selander *et al.*, 1986), were not involved in the establishment of the biofilm (Fig. 3 of I). Using indirect immunofluorescence, whole-cell ELISA, and Western blotting I showed that the average Mat fimbriae expression level in the cell population stayed constant throughout the transition from planktonic, free-floating phase to mature biofilm (Fig. 4 and 5 of I). This suggests that Mat fimbriae are not differentially expressed at certain stage of biofilm development.

Sequence comparison between the NMEC isolate IHE 3034 and the commensal K-12 strain MG1655 revealed that the predicted MatB-F proteins contains 27 amino acid differences in total (data not shown) that may affect the biogenesis and the functional properties of the fimbriae, as shown for type 1 fimbriae (Chen *et al.*, 2009; Pouttu *et al.*, 1999; Schembri & Klemm, 2001; Sokurenko *et al.*, 1998). To find out whether the identified Mat-mediated biofilm formation is specific for the NMEC host, and to assess if the allelic variation in Mat proteins affects biofilm-forming ability, I introduced expression plasmids encoding MatB-$F_{IHE\ 3034}$ and MatB-F_{MG1655}, into the wild-type MG1655 and the IHE 3034 *matBCDEF* mutant, and compared their effects on biofilm formation (Fig. 2 of I). Notable, ectopic expression of *matB-F* genes, derived either from IHE 3034 or MG1655, promoted biofilm growth in the non-NMEC host MG1655 at 20°C. This shows that Mat-mediated biofilm growth is independent on B2-phylogroup or NMEC -specific factors. Moreover, the two Mat fimbria variants did not differ in their capacity to support biofilm growth demonstrating that also M1655 harbours functional *matB-F* genes, and has thus retained the capability to express the fimbriae. However, while Mat fimbriae expression on the bacterial cell surface was observed at 20°C and 37°C after induction, biofilm was only produced at 20°C in both MG1655 and IHE 3034 *matBCDEF* backgrounds. This result illustrates that other temperature regulated factors than Mat fimbriae are critically involved in the biofilm process.

Temperature is a top-level cue for *E. coli* and other host-associated bacteria that signal the entry of the bacteria into host niche or the exit from a host into environment. The transition between ambient low temperature and body temperature have shown to result in dramatic, genome-wide changes in *E. coli* transcriptome, including differential expression of various genes associated with biofilm formation, outer membrane biogenesis, metabolism, and

transcriptional regulation as well as a large number of uncharacterized genes (White-Ziegler *et al.*, 2007; White-Ziegler *et al.*, 2008). One possibility for the the lack of Mat-fimbriae-dependent biofilm at 37°C is downregulation of other cell-surface associated structures interacting with the Mat fimbriae and/or the target surface. The second possibility for inability to form multicellular, sessile communities at 37°C despite induced Mat fimbriae expression is a presence of thermoregulated factors that mask or interfere with the function of short surface-located adhesins or signal receptors, and thus prevent biofilm formation by disturbing cell-surface initial interactions, close cellular contacts or cell-cell communication. Capsular polysaccharides constitute the outermost layer of the cell envelope and can act as a shielding factor of this nature (Schembri *et al.*, 2004). Further, certain capsulated *E. coli* strains can antagonize biofilm formation by a mechanism distinct from steric hindrance of surface adhesins. These strains, including IHE 3034, were shown to significantly release capsular antigens into the growth medium which inhibited biofilm development by weakening cell-surface contacts and cell–cell interactions (Valle *et al.*, 2006). The capsular K1 antigen, a 2,8-linked polysialic acid, is produced at host temperature but not at the low temperature of 18 or 22°C (Bortolussi *et al.*, 1983; Ørskov *et al.*, 1984). However, the laboratory strain MG1655 and other K-12 strains lack O-antigen (Stevenson *et al.*, 1994) and are also incapable of producing capsule. As Mat fimbriae-mediated biofilm formation was temperature-dependent also in this genetic background, the cell-surface associated or released K1 capsular polysaccharides cannot account the phenomenon.

Next, the role of the first gene of the operon, *matA*, in biofilm formation was investigated. *matA* is predicted to encode a protein of 196 amino acids that contains a putative, C-terminal helix-turn-helix (HTH) DNA-binding motif (Pouttu *et al.*, 2001; see Fig. 3A). The replacement of A166 in the scaffold helix (α2) and H179 in the recognition helix (α3) by proline was designed to disrupt the proper folding of the helices and consequently DNA binding (Barlow & Thornton, 1988). In both mutant strains, the proline substitution completely abolished surface expression of Mat fimbria (Fig. 3B) demonstrating the importance of functional MatA for Mat fimbriae expression. As expected from the previous results, the *matA* mutant lost its ability to biofilm growth, and the biofilm-forming phenotype was restored by single-copy chromosomal complementation of the mutation (Fig. 3C). Similar to *matB-F* genes, *in trans* overexpression of *matA* was not enough for biofilm formation at 37°C (Fig. 3C; Fig. 1 of III). A more detailed analysis of MatA and confirmation of its role as a regulator are provided in Chapters 4.2.2 and 4.2.6.

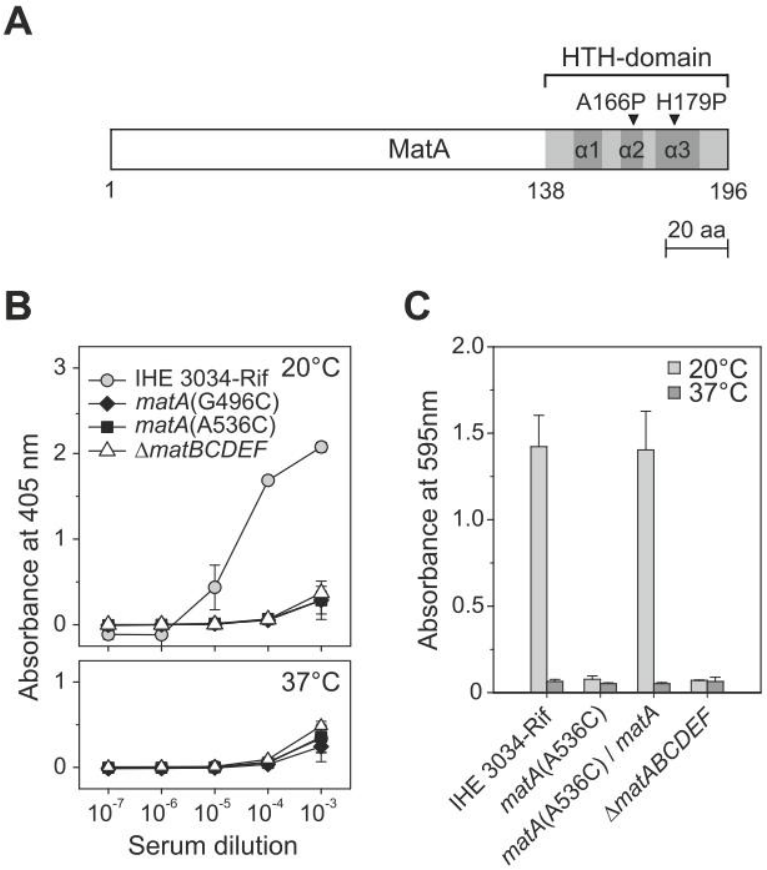

Figure 3. Proline substitutions at the HTH-domain of MatA abolish Mat fimbriae expression and biofilm formation in IHE 3034. (A) Schematic representation of MatA protein. The putative LuxR-type tri-helical helix-turn-helix (HTH) DNA-binding motif is indicated as a light grey box, and alpha helices of HTH are shown as dark grey boxes. The location of introduced A166P (resulting from G496C nucleotide substitution) and H179P (resulting from A536C nucleotide substitution) substitutions are indicated by arrowheads. (B) Surface expression of Mat fimbriae by the parent strain IHE 3034-Rif and *matA* mutants. Bacteria were analyzed by whole-cell ELISA with anti-Mat fimbria serum as primary antibodies after growth in LB medium at 20 and 37°C. The data represent means and standard deviations of two independent experiments. (C) Biofilm formation by IHE 3034-Rif derivatives after growth for 48 h in M63 medium in PVC microtitre plate wells at 20 or 37°C. The biofilm-associated cells were visualized by CV staining. The data shown were reproduced in three independent experiments, and error bars represent standard deviation. Assays with the Mat fimbriae-deficient IHE 3034 *matBCDEF* deletion mutant is shown for comparison.

4.1.3 The Rcs phosphorelay system is involved in Mat fimbriae-mediated biofilm formation (III)

Biofilm development is a complex, multifactorial genetic program, and to better understand the genetics of this process, IHE 3034 was subjected to mini-Tn5 transposon mutagenesis and the resulting library was screened for mutants exhibiting an altered biofilm phenotype in conditions supporting Mat-mediated biofilm formation. After testing of 4,418 individual transposon insertion mutants for their ability to develop biofilm on PVC, 32 candidate biofilm-forming mutants conferred a reproducible phenotype (Table 3 of III). Using direct DNA sequencing, the transposon insertions were mapped to 15 distinct genes, from which 11 were involved in lipopolysaccharide (LPS) biosynthesis. Many of these LPS-genes has

previously been linked to reduced biofilm capability in *E. coli* (Amini et al., 2009; Beloin et al., 2006; Niba et al., 2007; Puttamreddy et al., 2010). Next, we further screened the candidate mutants for general metabolic defects by measuring growth yield in M63 medium after biofilm growth and comparing with that of the parental strain. After this elimination, which included all LPS-defective mutants, eight mutants were retained. The expression status of the Mat fimbriae was tested by serum agglutination. Five of the mutants had lost both the ability to surface-express the fimbriae and to form biofilm: two independent insertions were identified in *matA*, and three insertions were in the coding region of *rcsB* gene at two different locations.

This result revealed a new regulatory cascade required for Mat fimbria expression as RcsB is a response regulator of the Rcs phosphorelay signalling pathway that responds to various, mainly stress-related, environmental cues (Clarke, 2010). Overexpression of *rcsB* from a multicopy plasmid restored Mat fimbria production and biofilm formation at 20°C, however, the strain did not gain the ability to display multicellular behaviour at 37°C (Fig. 1 of III). Further, when the high-copy-number plasmid encoding Mat fimbriae was transformed into *rcsB* mutant, the derivative behaved in manner similar to that of the wild-type strain. Given that RcsB is a pleiotropic regulator that controls various biofilm-associated genes (Ferrières & Clarke, 2003; Hagiwara *et al.*, 2003; Oshima *et al.*, 2002), bypassing the requirement of *rcsB* gene for biofilm formation is noteworthy. For instance, RcsB is absolutely required for colonic acid expression (Brill *et al.*, 1988; Gottesman *et al.*, 1985). This suggests that colonic acid is not compulsory for Mat-mediated biofilm formation and highlights the principal role of Mat fimbriae in the development of biofilm by IHE 3034.

One of the mutants with reduced biofilm capability possessed transposon within ECOK1_3592 ($tRNA^{Met}$) gene and showed moderate reduction in Mat fimbria expression. IHE 3034 carries ten loci for $tRNA^{Met}$, hence the defect in the mutant likely results in poor transcription of the adjecent genes: *argG*, which encodes the arginosuccinate synthase involved in the arginine biosynthetic pathway, and ECOK1_3591 encoding a conserved hypothetical protein (also known as RimP). In *E. coli* K-12, this $tRNA^{Met}$ locus, *metC*, forms multicistronic operon with *rimP-nusA-infB* genes (Régnier & Grunberg-Manago, 1989) and *rbfA-truB-rpsO-pnp* genes (Sands *et al.*, 1988). The genome of IHE 3034 harbours similar operon structures and thus the inserted mini-Tn5 may have a polar effect on these genes. Recently, screening of biofilm mutants in *Pseudomonas aeruginosa* PA14 identified over a half of these genes, namely *argG*, *nusA*, *rbfA*, *truB*, and *pnp*, in connection with reduced biofilm production (Müsken *et al.*, 2010).

The last two transposon mutants, *mprA* and *rfbA*, displayed enhanced biofilm phenotype. MprA (also known as EmrR) represents a third transcriptional regulator identified in this genetic screen. Northern blotting and serum agglutination confirmed that Mat fimbriae expression had remained intact in the *mprA* mutant (Fig. 2 and Table 3 of III). MprA represses the transcription of the chromosomal *emrRAB* multidrug resistance gene cluster, and the disruption of *mprA* leads to overexpression of the EmrAB pump (Lomovskaya *et al.*, 1995) which may explain the moderately enhanced biofilm phenotype of *mprA* transposon mutant. Efflux pumps are active in bacterial biofilms and highly upregulated during biofilm growth (Kvist *et al.*, 2008). Their inactivation, including EmrAB, has been shown to abolish

biofilm formation (Kvist *et al.*, 2008; Matsumura *et al.*, 2011). EmrAB also aids to pump out bile salts, detergents that disrupt biological membranes (Thanassi *et al.*, 1997). Recently, it was shown that the EmrAB pump has role in acid survival of *E. coli* (Deininger *et al.*, 2011), a mechanistic explanation however remains lacking.

The *rfbA* mutation prevents biosynthesis of dTDP-L-rhamnose, which is an important precursor of cell wall polysaccharides, e.g. the IHE 3034 O18ac antigen (Selander *et al.*, 1986) has a L-rhamnose residue in the linear backbone of the repeating tetramer unit (Gupta *et al.*, 1984). *rfbBDAC* gene cluster encodes the four enzymes commited to biosynthesis of dTDP-L-rhamnose. A glucose-1-phosphate thymidylyltransferase RfbA is responsible for the first enzymic step, which is the formation of dTDP-D-glucose from glucose-1-phosphate and dTTP (Samuel & Reeves, 2003). IHE 3034 harbours two genes for glucose-1-phosphate thymidylyltransferase, one in the *rfbBDAC* locus and the other clustered with a second *rfbB* gene and a *rfe–rff* enterobacterial common antigen locus. The two RfbA proteins share 66% amino acid identity. However, the phenotypic change of the *rfbA* transposon mutant indicates that the second *rbfA* gene cannot complement the mutation.

Taken together, our genetic screen for biofilm mutants was not saturating but we were able to identify several mutants affecting biofilm formation by IHE 3034. The most interesting gene identified as biofilm-important is *rcsB*, which encodes a response regulator of the Rcs phosphorelay singaling pathway. The loss of Mat fimbriation in *rcsB* mutant and the recovery of biofilm phenotype after ectopic expression of Mat fimbriae strongly indicate that RcsB affects the temperature-dependent biofilm formation by IHE 3034 primarily at the level of Mat fimbriae expression. Subsequent characterization of one of the three *rcsB* mutants is described in Chapters 4.2.2 and 4.2.4.

It should be noted that our mutagenesis screen and functional analysis of transposon mutants as well as type 1 fimbria, S fimbria, and flagellum -deficient derivatives were based on one-time point measuring the staining of the adherent biomass. This approach is very effective for assessing bacterial attachment and large-scale screening for mutants as it identifies the phenotypes that differ in terms of the extent of the biofilm formation. However, without performing the assay over several time points, it does not reveal the kinetics of attachment and biofilm development. The other drawback is that CV staining does not take the three-dimensional structure of the formed biofilm into account, e.g. flat and tightly packed structures or high and porous biofilms may bind equal amounts of stain and look similar in the spectrophotometric determination of CV. For instance, a colonic acid-deficient *E. coli* K-12 derivative had slower biofilm growth and formed more collapsed biofilm structure compared to wild-type but after a long-term incubation it closely resembled the wild-type strain in CV analysis (Danese *et al.*, 2000).

4.2 The regulatory network of Mat fimbriae expression and a motile-to-adhesive transition

4.2.1 The expression of fimbrillin *matB* mRNA is dictated by temperature-responsive *mat* promoter and involves mRNA processing (II)

To characterize the expression of the *mat* operon in IHE 3034, I first performed Northern blot analysis with probes specific for each *mat* gene (Fig. 3 of II). Only a probe complementary to the major subunit gene *matB* detect a monocistronic transcript, 0.67 kb in length, in the RNA harvested from logarithmic-phase culture of IHE 3034-cells grown at 20°C. Subsequent analysis of cotranscription by RT-PCR revealed that one large polycistronic transcript, *matABCDEF*, is transcribed from the *mat* gene cluster (Fig. 3 of II). In line with this result, multicopy *lacZ* reporter analyses of promoter activities driven by upstream region of each *mat* gene showed that the transcriptional fusion *pA* encompassing the *matA* upstream region exhibited a high level of β-galactosidase activity, whereas the expression of the other reporter fusions, *pB* to *pF*, were at the level of the vector control (Fig. 4 of II). Thus, we were unable to detect additional regulatory regions that would have been active under the test conditions. Quantification of the transciption level of the individual *mat* genes by real-time RT-PCR revealed that the transcription of the *mat* operon is highly polar (Fig. 3 of II) suggesting that differential gene expression for the production of Mat fimbriae is achieved through partial termination, processing and/or differential stability of mRNA segments of the polycistronic transcript. The sequence immediately downstream of the *matB* stop codon contains two predicted stem-loop structures. The latter putative secondary structure has features of an intrinsic, Rho-independent transcriptional terminator (d'Aubenton Carafa *et al.*, 1990), which suggests that transcription of *matB* can be terminated immediately downstream of the open reading frame.

The results above indicate that the *matB* mRNA is not transcribed from a separate promoter but rather is processed from *matAB* and/or *matABCDEF* precursors. The processing site was mapped to the *matA-matB* intergenic region by primer extension (Fig. S2 of II). RNase E is a principal endonuclease in *E. coli* and involved in regulation of several fimbrial operons (Balsalobre *et al.*, 2003; Nilsson & Uhlin, 1991; Nilsson *et al.*, 1996). To test whether it is also recruited in the maturation of the *matB* transcript, I introduced the silencing *rne131* mutation (Kido *et al.*, 1996; Lopez *et al.*, 1999) in the genome of IHE 3034 and evaluated its effect on *matB* mRNA profile. The RNase E-deficient cells had higher intracellular levels of *matB* transcripts, including higher molecular weight transcripts than the monocistronic *matB* mRNA (Fig. 4). However, the major 0.67-kb mRNA is the same species as observed for IHE 3034 wt. This indicates that the maturation of the *matB* transcript is independent of Rnase E, as in the case of F1845 and CFA/I fimbrillin mRNAs (Bilge *et al.*, 1993; Jordi *et al.*, 1993). The roles of other endoribonucleases, such as RNase III, RNase G, or RNase P (Deutscher, 2006), in the maturation of the *matB* transcript remain open at the present.

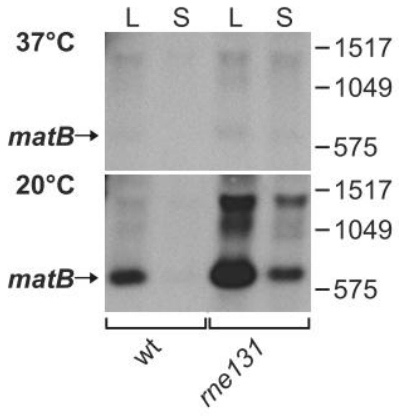

Figure 4. Processed *matB* mRNA accumulates in *rne131* mutant. Northern blot analysis of RNA harvested from wild-type IHE 3034 and *rne131* mutant strain that produces truncated, nonfunctional RNase E lacking the noncatalytic C-terminal region. Northern blots were performed with *matB*-probe on 2.5 µg of total RNA isolated from cells grown in LB to mid-logarithmic (L) and stationary (S) phases at 20°C and 37°C. The position of the monocistronic *matB* message is indicated with arrows. The lenght (bp) of DIG-labelled RNA markers are shown on the right for reference.

The expression of the transcriptional fusion *pA-lacZ* was stronger at 20°C than at 37°C (Fig. 4 of II), which agrees with the Nothern blotting analysis of *matB* mRNA (Fig. 4; Fig. 3 of II). The temperature sensitivity was not significantly affected by the previously observed temperature response of the weak transcriptional terminator *trpt* in the *lacZ* leader sequence (Liang *et al.*, 1998; Repoila & Gottesman, 2001). Elucidation of the transcriptional initiation sites identified three different 5′-termini upstream of *matA* (see Fig. 5), from which the second (T2) transcriptional start site is preceded by strong -35 and -10 promoter motifs (Fig. S4 of II) and was associated with the highest transcriptional initiation activity (Fig. 4 of II).

4.2.2 MatA and RcsB cooperatively activate the *mat* promoter (II, III)

The lack of N-terminal signal peptide and the presence of functional, tri-helical HTH domain in the predicted MatA protein as well as Mat-fimbriae-deficient phenotype of *matA* HTH-mutants (see Chapter 4.1.2) strongly indicate that MatA is a non-structural protein that has a regulatory function in *mat* expression. Northern blotting demonstrated that IHE 3034 *matA* mutants had completely lost capacity to accumulate *matB* mRNA (Fig. 6 of II). Chromosomal single-copy complementation of the *matA* substitution restored the the level of *matB* mRNA to that of wild-type whereas a constitutive, ectopic expression of $MatA_{IHE\ 3034}$ in the mutant strain resulted in a highly elevated accumulation of *matB* mRNA exceeding of the temperature and growth phase constraints. Also $MatA_{MG1655}$ variant, which differs in six residues compared to $MatA_{IHE\ 3034}$, complemented the mutation but had lower capacity to enhance *matB* transcription at 37°C. In MG1655, overexpression of *matA* did not activate *matB* expression.

The Mat-fimbria- and biofilm-defective *rcsB* transposon mutant acted as *matA* HTH-mutants in Northern blotting analysis as no *matB* mRNA was detected in the RNA harvested from *rcsB* cells grown at 20°C (Fig. 2 of III). Complementation of *rcsB* using multicopy plasmid encoding wild-type allele (identical between IHE 3034 and MG1655) revealed that the *trans*-regulatory activity of RcsB differs from that of MatA; *in trans* overexpression of RcsB restored the transcription of *matB* to, but not beyond, normal levels. Subsequent complementation analysis demonstrated that neither MatA nor RcsB protein can functionally substitute for the other supporting their critical role in *mat* expression.

For monitoring the influence of MatA and RcsB on transcription initiation from the *mat* promoter, I inserted two putative regulatory sequences, *pA* and *pAB*, downstream of transcriptional terminators and upstream of the promoterless *lacZ* gene, and integrated these constructs into *lac* locus of *matA*(A536C) and *rscB*::Tn5 mutants as well as their parental strains. Both *pA* and *pAB* fragment encompasses 0.6 kb of 5' upstream region of *matA* containing three transcription start points. The longer *pAB* further extends beyond the *matA* coding region, designed to mimick the early-stage of *matB* transcription (see Fig. 5). Both *pA*- and *pAB-lacZ* promoter fusions showed significant temperature-responsive transcription initiation (p <0.005) in the wild-type background, and comparison of their expression profiles indicated the presence of downstream regulatory element (DRE) lying downstream of *mat* promoter (Fig. 7 of II; Fig. 3 of III). The DRE has an effect of reducing the level of transcription that reaches *matB*. This putative DRE exerted a repressive effect that was notably stronger at nonpermissive temperature of 37°C than at 20°C resulting only basal levels of expression from the *pAB-lacZ* promoter fusion at 37°C. β-galactosidase activity measurements directed by the *pAB-lacZ* fusion demonstrated that no induction of transcription occurred at 20 and 37°C in cells lacking a functional MatA or RcsB. In

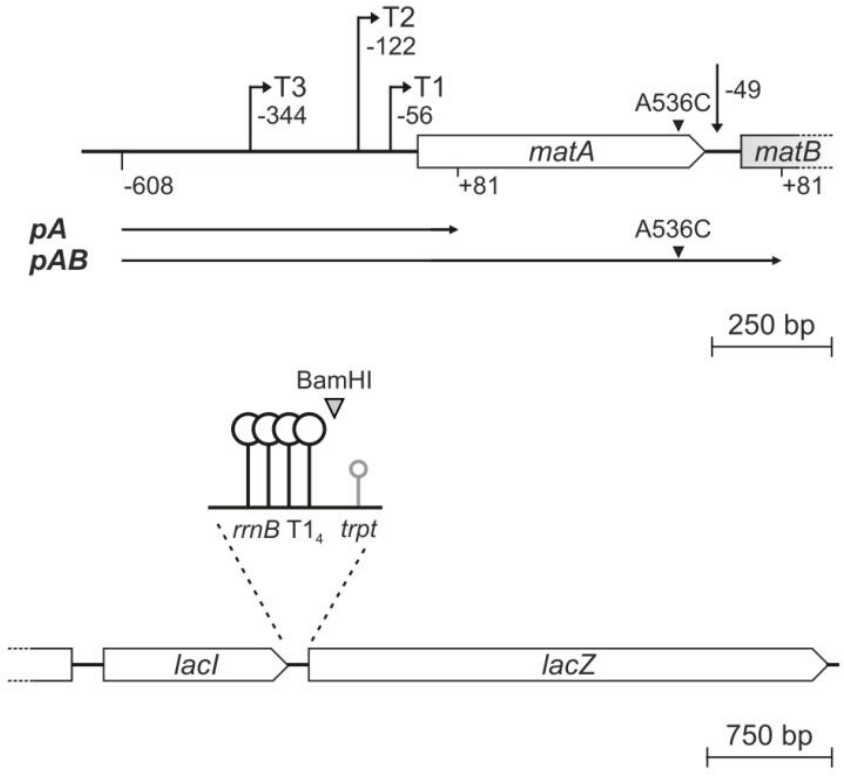

Figure 5. An outline of the IHE 3034 *mat* DNA fragments *pA* and *pAB* used to construct single-copy transcriptional *lacZ* fusions. Right-angled arrows indicate the three transcription start sites, and a vertical arrow shows the 5' end of monocistronic *matB* mRNA identified in this study. The numbers in the upper scheme correspond to the distances from the translational start of *matA* or *matB*. The location of introduced A536C missense mutation in chromosomal *matA* and *matA* in the *pAB* fragment is indicated by a black triangle in the upper scheme. The lower scheme shows a promoterless *lacZ* reporter at the *lac* locus in the chromosome of IHE 3034. The native *lacI-lacZ* intergenic region was replaced with a fragment of pRS551 that contains a transcription termination region with four *rrnB* T1 terminators (stem-loop symbols), and a promoter cloning site BamHI (grey triangle). The weak *trpt* terminator in the *lacZ* leader sequence is indicated as a grey stem-loop.

contrast, deletion of genes known to be involved in the modulation of the phosphorylation state of the transcription factor RcsB, namely *rcsC*, *rscD* as well as the *ackA-pta* genes (Fredericks *et al.*, 2006; Schmöe *et al.*, 2011; Takeda *et al.*, 2001), did not decrease the *pAB* reporter expression. These results indicate that the RcsB activity is controlled by an alternative unknown phosphoryl donor, as proposed by others (Fredericks *et al.*, 2006; Majdalani & Gottesman, 2005; Majdalani *et al.*, 2005), or the activation of *mat* promoter is independent of RcsB phosphorylation and instead the activity of RcsB is modulated by interaction with other transcriptional regulators. Consistent with Northern blot results (Fig. 6 of II; Fig. 2 of III), plasmid-derived RcsB was unable to enhance the promoter activity over normal levels whereas overexpression of *matA* caused high upregulation of promoter expression at both temperatures (Fig. 7 of II; Fig. 3 of III). Thus, MatA autoregulates its own expression through a positive feedback loop that necessitates the presence of RcsB response regulator. The connection between MatA and RcsB was shown to be unidirectional as MatA had no regulatory activity on *rcsDB* and *rcsB* transcription initiation or accumulation of *rcsB* mRNA (Fig. 3 of III).

4.2.3 The *mat* operon is under the repressive action of H-NS (II)

Ectopic expression of *matA* bypassed the temperature control of the *mat* promoter, which suggests that activation by MatA results from competition with regulator(s) that repress(es) transcription at 37°C. Employing mini-Tn5 transposon mutagenesis and enrichement with immunomagnetic particles, we identified the nucleoid-associated protein H-NS, a well-characterized pleiotropic regulator of *E. coli* (Dorman, 2004; Stoebel *et al.*, 2008), as a repressor of *mat* expression in IHE 3034. The H-NS deficiency elicited upregulation of all the six *mat* genes at both temperatures (Fig. 10 and S6 of II). The *matB* transcription and the *mat* promoter activity in H-NS-deficient cells resembled the patterns seen in the *matA* overexpression hns^+ derivative while the functional inactivation of MatA by the H179P substitution in the *hns* background decreased the transcription initiation and the mRNA accumulation (Fig. 11 of II). The expression profiles of single-copy *pA*- and *pAB-lacZ* promoter fusions in *hns* mutant indicated that the temperature-responsive negative effect of DRE is largely H-NS-mediated. It has been previously shown that DREs located in *eltAB*, *kpsMT*, and *proU* loci are crucial for temperature-dependent repression by H-NS (Bouffartigues *et al.*, 2007; Xue *et al.*, 2009; Yang *et al.*, 2005). Together, the results of this thesis demonstrate that MatA acts as an antirepressor overcoming the blockade exerted by H-NS but also has an additive effect on the *mat* promoter activation.

As descibed in Chapter 4.1.2, the K-12 strain MG1655 has presumably a cryptic but functional *mat* operon. The finding of H-NS prompted us to test whether this global repressor silences the *mat* expression also in K-12. The inactivation of *hns* in MG1655 abrogated the tight repression of *mat* operon and enabled the surface expression of Mat fimbriae at both 20 and 37°C. However, the induction of *matB* transcription was notably stronger at 20°C as compared to the transcription at 37°C (Fig. 10 of II), and the reason for this was traced to the less effective $MatA_{MG1655}$ variant that is unable to further enhance *matB* expression at 37°C in the MG1655 *hns* mutant.

4.2.4 MatA, RcsB and H-NS as well as CRP-cAMP regulators directly interact with the *mat* regulatory DNA (II, III and this thesis)

The identified regulators of *mat* operon, MatA, RcsB, and H-NS, can exert their regulatory function by direct binding to a *mat* promoter or indirectly by affecting expression of another regulatory gene(s) which, in turn, affect(s) expression of the *mat* operon. To determine whether MatA, RcsB, and H-NS affect *mat* expression by a specific interaction with the *mat* regulatory region, the proteins were purified for electrophoretic mobility shift assays (EMSAs). Our numerous attemps to produce purified, soluble recombinant N- or C-terminally histidine-tagged MatA or GST-MatA failed. However, the corresponding genes complemented the *matA*(A536C) mutation *in vivo* indicating that a small fraction of the recombinant proteins were properly folded (data not shown). To improve the solubility of MatA, the protein was expressed as a fusion protein carrying at its N-terminus the effective solubility-enhancing 44-kDa maltose-binding protein (MBP) (Kapust & Waugh, 1999). The majority of MBP-MatA was present in an insoluble form in the cells, but extensive optimization of expression conditions resulted in production of soluble fusion protein in quantities enough for functional studies. The chaperonin GroEL involved in the correct folding of nascent proteins (Buchberger *et al.*, 1996) co-purified at equal relative amounts along with the His_6-, GST-. and MBP-MatA preparations (Fig. S10 of II; data not shown). MBP-RcsA and His-tagged RcsB, a constitutively active $RcsB_{D56E}$ (Gupte *et al.*, 1997), and H-NS were purified without problems.

EMSA demonstrated that MBP-MatA bound in a concentration-dependent and a temperature-independent manner to the *pmatA* fragment, containing the primary transcriptional start site T2 upstream of *matA* (see Fig. 6A), and the binding was diminished by proline substitution at the HTH-domain (H179P) of MBP-MatA (Fig. 12 and S7 of II). The *pmatA* promoter fragment was also bound by $RcsB_{D56E}$, RcsB, and H-NS but not by RcsA or MBP (Fig. 12 and S7 of II; Fig. 4 of III; Fig. 4 of IV). The binding of $RcsB_{D56E}$ and RcsB to the *mat* regulatory DNA presumably involves a potential 14-bp RcsAB box upstream of the transcriptional start site T2 (see Fig. 6A). In line with this, site-specific mutations of conserved nucleotides in the RcsAB box resulted in total loss of *mat* promoter expression demostrating the importance of the box for initiation of *mat* transcription (Fig. 4 of IV). None of the purified proteins showed efficient shifts with a *pmatB* fragment harbouring sequences surrounding the *matA-matB* intergenic region and the 5' processing site of *matB* mRNA. These findings are in agreement with the multi- as well as single-copy *lacZ* reporter analyses, which indicated that the *matB* mRNA is not transcribed from a separate promoter but is rather originating from the main promoter upstream of the operon.

Two high-throughput analyses of distribution of binding sites for CRP along the *E. coli* genome, utilizing method of ChIP-chip (chromatin immunoprecipitation combined with microarray hybridization) and an improved genomic SELEX (systematic evolution of ligands by exponential enrichment), identified CRP-binding sites in the *mat* operon, one located upstream of *matA* (Grainger *et al.*, 2005) and the other within *matB* coding region (Shimada *et al.*, 2011) (see Fig. 6A). These results suggest that the signal transduction system CRP-cAMP, in addition to Rcs system, might have a role in the regulation of Mat fimbriae expression. Previous studies have demonstrated that multiple chaperone-usher

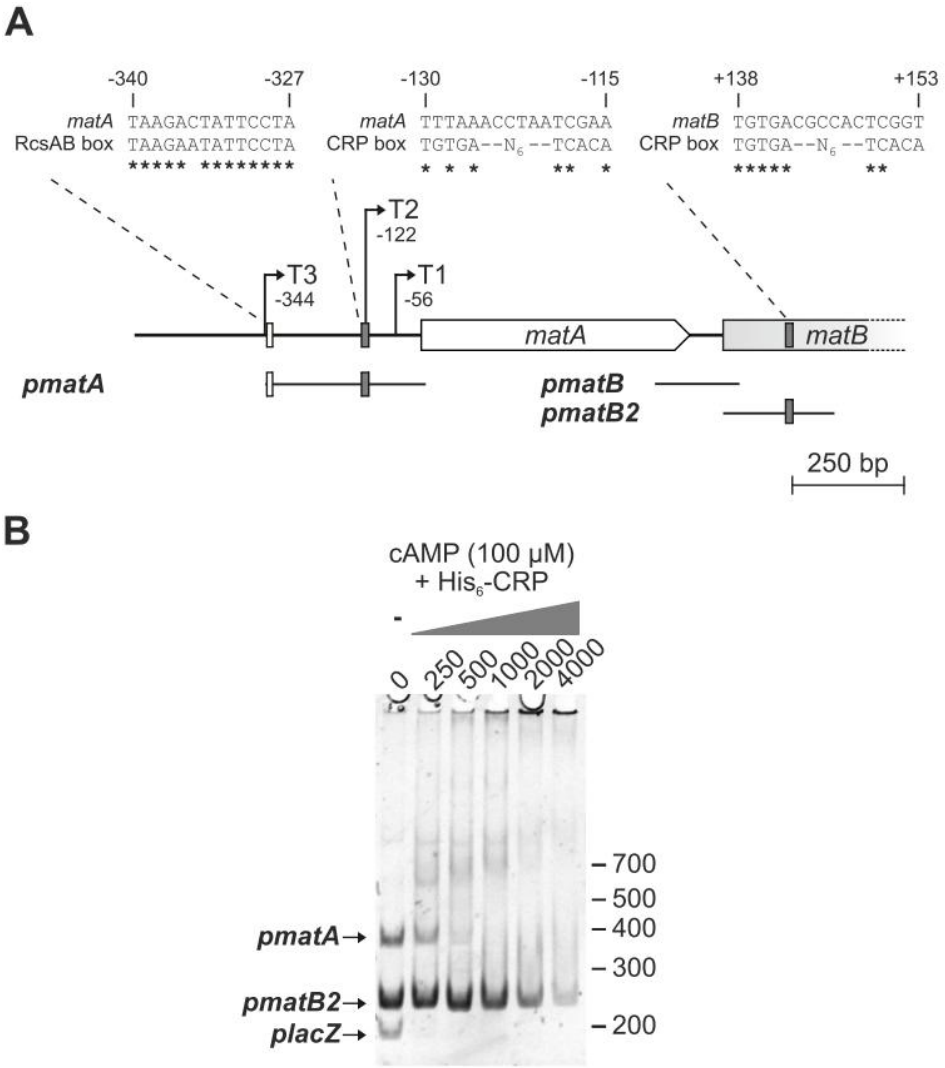

Figure 6. Binding of cAMP-CRP to the *mat* upstream regulatory region and within the *matB* coding region. (A) Schematic presentation of the *matAB* region of IHE 3034 and DNA fragments used in EMSA. The three transcriptional start sites upstream of *matA* (T1 to T3) are indicated with right-angled arrows. The positions of putative RcsAB and CRP boxes are shown as white and grey boxes, respectively, and aligment results of the *matA* upstream region with the consensus sequences of RcsAB box (Wehland & Bernhard, 2000) or CRP box (Berg & von Hippel, 1988; de Crombrugghe *et al.*, 1984) are indicated. Numbering denotes the distances from the translational start site of *matA* or *matB*. (B) Competitive EMSA performed in the presence of *pmatA* and *pmatB2* fragments with increasing concentrations (nM) of His_6-CRP proteins together with 0.1 mM cAMP. The reaction mixtures were incubated and separated by electrophoresis at 20°C, and DNA fragments were visualized by ethidium bromide staining. The length of DNA standards in bp are shown on the right for reference.

fimbriae are regulated by CRP-cAMP regulatory complex (Båga *et al.*, 1985; Edwards & Schifferli, 1997; Korea *et al.*, 2010; Müller *et al.*, 2009; Weyand *et al.*, 2001). To verify the interactions of CRP-cAMP regulatory complex with *mat* operon, we purified His-tagged *E. coli* IHE 3034 CRP under native conditions and measured the direct binding of CRP to the *pmatA* and *pmatB2* fragments containing the proposed binding sites and to the positive control fragment (*placZ*) in the presence of cAMP (Fig. 6B). EMSA depicted that the cAMP-CRP complex interacted with the regulatory region upstream of *matA* in a CRP dose-

dependent manner, although with lower affinity compared to the *lacZ* promoter, while much higher concentrations (2 μM) were needed to alter the migration of the *pmatB2* fragment. The CRP-cAMP binding to the *mat* upstream region prompted us to construct a chromosomal deletion of *crp* gene in IHE 3034 using the method of Datsenko and Wanner (2000). However, we did not detect noticeable changes in the surface expression of Mat fimbriae in IHE 3034 *crp* mutant cells with and without a *crp* complementation plasmid grown in LB (assayed by ELISA) or M9 minimal media supplemented with glycerol or glucose (assayed by serum agglutination) compared to the parental wild type cells (data not shown). Thus, the *in vivo* role of low-affinity binding of CRP to the *mat* regulatory region remains to be elucidated.

4.2.5 Post-transcriptional control of *mat* operon involves MatA-mediated stabilization of *matB* mRNA (II)

The regulatory arrangements driving the expression of Mat fimbriae involve post-transcriptional regulation for the control of *matB* transcript levels. Our results showed that the genes of the *mat* operon are sequentially transcribed from a single promoter, and the resulting polycistronic mRNA(s) is subject to a rapid processing event that results in the accumulation of a *matB*-coding mRNA, similar to the generation of the major fimbrillin transcripts of *pap*, *sfa*, *cfa*, and *daa* operons (Balsalobre *et al.*, 2003; Bilge *et al.*, 1993; Båga *et al.*, 1988; Jordi *et al.*, 1993). The detection of high amounts of *matB* transcript but not other *mat*-specific mRNAs suggests segmental differences in stability within the polycistronic transcript. In *E. coli*, most mRNAs are rapidly degraded, and the process is largely triggered by RNase E-mediated endonucleolytic cleavages, followed by a rapid 3′ digestion by exoribonucleases, and a final finishing to mononucleotides by oligoribonuclease (Arraiano *et al.*, 2010). Evaluation of *matB* mRNA profile of the *rne131* mutant and an isogenic rne^{+} strain IHE 3034 by Northern blot analysis (Fig. 4) revealed accumulation of *matB* mRNA species in the RNase E-deficient cells indicating a role for RNase E in the stability of *matB* mRNA.

mRNA decay assays demonstrated that the stability of the monocistronic *matB* mRNA is strongly affected by temperature; the half-life was a sixfold greater at 20°C compared to that of 37°C (Fig. 8 and S3 of II). At low temperature of 20°C, the *matB* transcript is unusually stable with a mean half-life exceeding 38 min in exponentially growing IHE 3034 cells. Despite the enhanced *matB* mRNA steady-state levels in the *rne131* mutant, the loss of functional RNase E delayed the turnover of the corresponding transcript only at 37°C (by 2.6-fold to 17.6±1.6 min, $p < 0.002$, data not shown).

Sequences immediately downstream of the *matB* open reading frame contains two regions of dyad symmetry that would allow the formation of two stem-loop structures at the 3' end of the *matB* mRNA. Such secondary structures have been shown to act as protective barriers against 3′ exoribonucleolytic degradation (Blum *et al.*, 1999). The processes of bacterial transcription and translation are coupled events, and translation can initiate soon after the ribosome-binding site (RBS) has been transcribed. Association of mRNA with translating ribosomes can greatly affect the longevity of individual mRNA, as endonucleolytic cleavage sites can be occluded by ribosomes (Deana & Belasco, 2005). Thus, tempereture-dependent

formation of secondary structures and/or enhanced translation during growth at low temperature, as seen for *rpoS* (Sledjeski *et al.*, 1996), may account for the longevity of the *matB* transcript at 20°C. In contrast, at 37°C the monocistronic *matB* mRNA seems to be less protected from ribonuclease attack and is exposed to degradation initiated by RNase E.

The mechanism for the differential stability of *matB* mRNA may involve a binding of a regulatory protein or a small untranslated RNA (sRNA) to the transcript. It has been shown that the RNA-binding factors can affect the decay rate of transcripts by e.g. accelerating directly RNase E–dependent degradation (Pfeiffer *et al.*, 2009), or by affecting translational efficiency through a decrease in ribosome binding (Bouvier *et al.*, 2008; Darfeuille *et al.*, 2007) or disruption of an inhibitory secondary structure, that would otherwise sequester the ribosome-binding site (Majdalani *et al.*, 1998; Majdalani *et al.*, 2002). I showed that the MatA activator protein not only acts at the promoter level but also contributes to the post-transcriptional regulation of *mat* operon. Ectopic overexpression of *matA* significantly increased the half-life of the processed *matB* mRNA both at 20 and 37°C (Fig. 8 of II). While the mean half-live of the *matB* transcript at 37°C was 6.7-6.9 min in IHE 3034 cells with and without the vector pACYC184, it increased about 4-fold to 29±3 min in the presence of plasmid carrying constitutively expressed *matA*. Closely resembling improvement of *matB* mRNA stability was observed in *hns* mutant (Fig. 11 of II). Since we detected an elevated *matA* transcript level in the H-NS deficient strain (see Chapter 4.2.3), the *matB* half-life profile of the *hns matA* double mutant was also determined. The loss of functional MatA had no effect on the stability of *matB* transcript in *hns*-deficient cells growing at 20°C, and at 37°C the half-life partially decreased but was still significantly higher (4-fold) than what was observed in the wild-type cells. These results suggest that MatA is not exerting its stabilizing effect on the *matB* mRNA through direct binding but rather indirectly via another regulatory protein or sRNA that stabilizes the transcript. In this model, H-NS decreases *matB* mRNA stability by repressing the expression of the RNA-binding factor, and MatA is able to counteract this repression. When *matA* is overexpressed, in *hns* mutant or ectopically expressed, the expression of the RNA-binding factor is further enhanced. This is due to an additive effect of MatA in the activation of transcription initiation, either by alleviating additional repression mediated by negative regulatory factors other than H-NS or by acting as a classical transcriptional activator that improves the recruitment of the RNA polymerase to the promoter.

An alternative explanation for our observations of increased *matB* stability in *hns* mutant and in the presence of multicopy *matA* is that there is too much *matB* transcripts (due to opened promoter) to be efficiently degraded, a secondary consequence of exceeding the capacity of RNase E. However, if this is the full explanation of the *matB* half-live profiles in these strains, we would expect similar profile in *rne131* mutant. This was not the case; the loss of functional RNase E did not increase the half-life of *matB* as much as loss of H-NS or overproduction of MatA. However, there could be other endoribonucleases, such as RNase III, involved in the degradation of *matB* transcript so we cannot completely exclude the hypothesis.

4.2.6 MatA inversely coordinates two opposite functions, Mat fimbriae-mediated adherence and flagella-driven motility (IV)

During phenotypic testing of IHE 3034 *matA* derivatives, we noticed by phase-contrast microscopy that constitutive ectopic expression of *matA* rendered the bacterium nonmotile indicating that the expression of regulator MatA is linked to a flagellar system malfunction. Examination of swimming motility phenotypes showed that *matA*(A536C) mutation had a minor positive effect on motility at 20°C both in IHE 3034 and MG1655 backgrounds, whereas complementation with pACYC184-derivatives carrying the *matA* gene from IHE 3034 or from MG1655 severely impaired the motility at 20°C as well as at 37°C (Fig 1 of IV). In contrast, the co-expression of *matBCDEF* from the plasmid or deletion of chromosomal *matBCDEF* genes did not drastically affect the motility of IHE 3034.

Northern blot and single-copy *lacZ* reporter analyses revealed that MatA controls motility at the top level of a tiered flagellar regulatory cascade by repressing the expression of the flagellar master operon *flhDC* (Fig. 2 and 3 of IV). In line with the motility phenotype, the inactivation of *matA* elicited moderately upregulated transcription initiation from the class 1 *flhDC* promoter resulting increase in *flhDC* transcription, while *in trans* overexpression of MatA severely impaired the ability to activate the promoter and consequently resulted in total loss of *flhDC* mRNA. The expression pattern of flagellin gene *fliC* at the distal end of the cascade correlated that of *flhDC*. EMSA assays demonstrated a direct interaction of MatA with the upstream region of the *flhDC* operon; the binding affinity was similar to that observed with the *mat* regulatory region (Fig. 4 of IV).

Taken together, the results demonstrate that MatA can act beyond its local influence on the *mat* operon as an additional component of the complex regulatory network controlling flagellar expression. The cross-talk with flagella system is biologically important as it enables an inverse coordination of largely antagonistic forces, flagella-driven locomotion and fimbriae-mediated immobilization. Flagellar motility facilitates approach of a suitable surface and helps to overcome surface repulsion, and the subsequent cell-surface adhesion or cell-cell interaction is commonly mediated by fimbrial structures. Upon bacterial attachment to a surface, force-generating flagellar movement can destabilize the contact to the surface. Cleary and co-workers (2004) showed that flagellar synthesis of enteropathogenic *E. coli* strain E2348/69 greatly decreased in the course of microcolony formation on epithelial cells, and another study found that the epithelial cell adhesion of this strain led to upregulation of Mat fimbriae (Saldaña *et al.*, 2009). Our results indicate that these two phenomena are connected by the opposite regulatory actions of MatA.

To my knowledge, MatA is the first master fimbriae regulator that represses the transcription of *flhDC* operon in *E. coli*. The flagellar inhibition by the curli activator CsgD occurs at a step after class 1 transcription by repression of *fliE* and *fliEFGH* operons (Ogasawara *et al.*, 2011). Overexpression of P fimbriae protein PapX in UPEC (Simms & Mobley, 2008), S fimbriae protein SfaX in NMEC (Sjöström *et al.*, 2009a), and MrpJ in *Proteus mirabilis* (Li *et al.*, 2001; Pearson & Mobley, 2008) down-regulate motility but have no impact on the fimbriae expression. Moreover, a recent study of Reiss and Mobley (2011) indicated that the

effect of *papX* on motility involves pathogen-spesific host factors as the K-12 strain MG1655 remained motile when harbouring multicopy *papX* although the strain has an intact binding site for PapX in the *flhD* promoter. In this thesis work, however, motility was repressed in both the pathogenic strain IHE 3034 and the non-pathogenic strain MG1655 when *matA* was ectopically expressed (Fig. 1 of IV).

It can be argued that the effect of MatA on motility is really seen only when the regulator is constitutively expressed which challenges the physiological role of this regulatory circuit. However, it should noted that the experiments in this thesis study describe the average behavior of the cells in a population and do not take into account of cell-to-cell variability. Mat fimbriae are synthesized in a phase-variable manner (Pouttu *et al.*, 2001; Fig. 4 of I; see Fig. 7), generating phenotypic heterogeneity within the population, which can mask the motility effect at a population level. Approximately 10% of the cells applied for Northern blotting and *lacZ* reporter experiments express Mat fimbriae at 20°C (Pouttu *et al.*, 2001). According to Chapter 4.2.2, MatA forms a positive feedback loop in which it reinforces its own expression, and in turn enables the expression of the fimbria. Immunofluorescence microscopy analysis indicated that the variation of Mat fimbriation is significantly influenced by the concentration of MatA, as ectopic expression of *matA* resulted in a uniform Mat fimbriation "on" status in the culture (Fig. 7). These results are compatible with the model where the variability is epigenetic, in which a change in gene expression does not involve changes in DNA sequence, and contains a stochastic, reversible switch. The switch depends on the presence of positive feedback loop and determines the level of the key regulator MatA. Such a system can generate a bistable output, i.e. two alternative expression

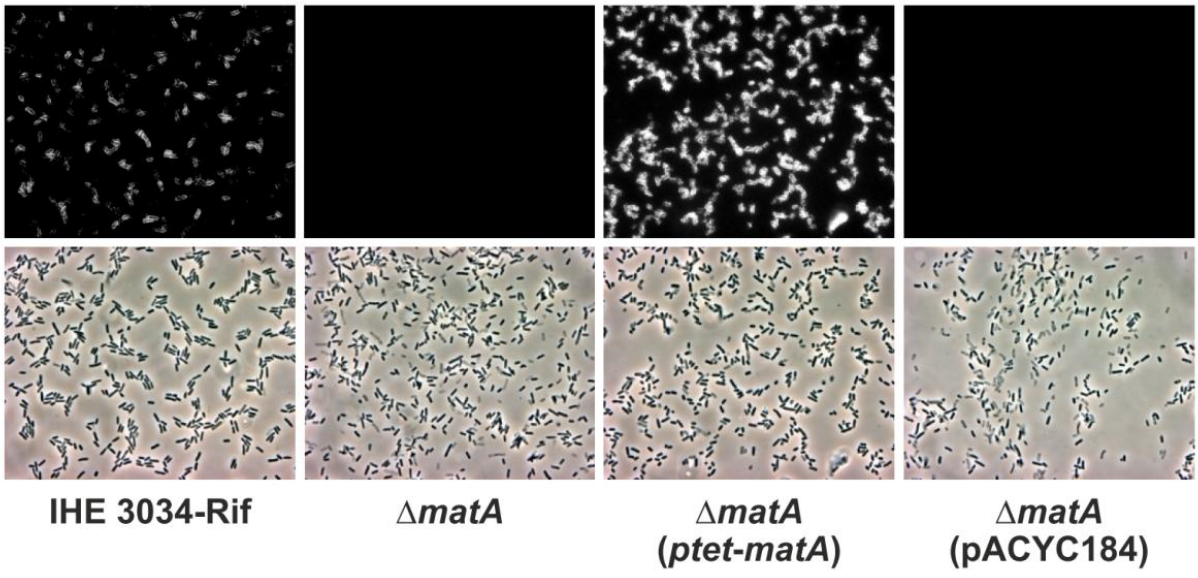

Figure 7. Indirect immunofluorescence microscopy analysis of the phase-variable Mat fimbriae.The parental strain IHE 3034-Rif and the *matA deletion* derivatives were grown in LB at 20°C, and Mat fimbriae were detected using anti-Mat fimbria antiserum as primary antibodies. The *matA* deletion was complemented with a pACYC184-based plasmid carrying the *matA* gene under the constitutive *ptet* promoter of the vector. Representative microscopy fields of epifluorescence (upper panels) and phase-contrast (lower panels) images are shown.

states among cells within an isogenic population (Dubnau & Losick, 2006; Smits *et al.*, 2006). This model may explain the consistent but minor effect of *matA* mutation on motility and *flhDC* promoter activity. More importantly, this model predicts that individual cells expressing Mat fimbriae are less motile, at least if the threshold levels for *matA* autoactivation and *flhDC* repression are similar. Thus, we hypothesize that the MatA-*flhDC* cross-talk plays a role in a transition from a motile free-living phase to a Mat fimbria-mediated sessile growth.

The *flhDC* connection incorporates MatA tigtly to the complex flagellar regulatory network via two feedforward loop circuits, RcsB-MatA-*flhDC* and H-NS-MatA-*flhDC* (Fig. 8). The effect of the abundant, pleiotrophic regulator H-NS on motility is opposite to that of MatA, as it activates the expression of the *flhDC* operon (Bertin *et al.*, 1994) by binding to the regulatory region of the operon (Soutourina *et al.*, 1999). In consequence, an *hns* mutant is nonmotile. As described in Chapter 4.2.3, H-NS represses the transcription of the *mat* operon, including the *matA* gene. As loss of H-NS causes upregulation of *matA*, the motility defect of an *hns* mutant may partially be due to the MatA-mediated repression of *flhDC*. However, the *hns matA* double mutant is still nonmotile (data not shown), which indicates that the activating effect of H-NS on *flhDC* is stronger than the repressive effect of MatA. Given that H-NS protein is also required for the normal flagellar motor function (Ko & Park, 2000; Paul *et al.*, 2011), MatA cannot solely be responsible for the locomotion deficiency of the *hns* mutant. However, the link between MatA, H-NS, and FlhDC is noteworthy as this circuit fits to a regulatory network motif termed the cohorent type-4 feedforward loop (Mangan & Alon, 2003; Fig. 7). The feedforward loop (FFL) is one of the most prevalent network motifs in transcription networks. In the FFL, a transcription factor *X* regulates the promoter of the gene for *Y*, and both *X* and *Y* regulate the promoter of target gene *Z*. Thus, *Z* is regulated by a direct path from *X* and by an indirect path through *Y*. H-NS is able to promote motility directly via binding to the *flhDC* promoter and indirectly by repressing the transcription of *matA*. Furthermore, MatA forms together with the Rcs phosphorelay regulator RcsB another cohorent feedforward loop (type-3) that influences motility (Fig. 8). RcsB has an opposite regulatory effect on the expression on *flhDC* and *mat* operon; it serves as a repressor acting in the *flhDC* promoter region (Francez-Charlot *et al.*, 2003) and as an activator for *mat*. Thus, MatA and RcsB jointly inhibit flagella synthesis and promote the surface attached mode of growth.

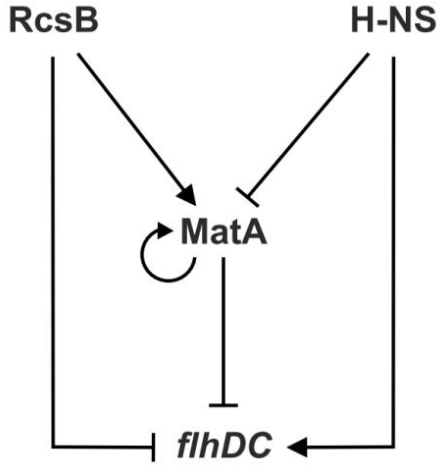

Figure 8. Autoregulated MatA is integrated into two feedforward loop (FFL) network motifs that control the expression of the flagellar master operon *flhDC*. The RcsB-MatA-*flhDC* circuit is in the form of a cohorent type-4 FFL, and the H-NS-MatA-*flhDC* circuit in the form of a cohorent type-3 FFL. Arrowheads indicate activation while blunt ends symbolize repression.

4.2.7 Host-associated cues, acidic pH and high acetate concentration, promote Mat fimbriae expression at 37°C in *E. coli* isolates of the B2 phylogenetic group (II)

Low temperature is the only environmental condition that was previously established to stimulate Mat fimbriae expression in IHE 3034 (Pouttu *et al.*, 2001). Girón's laboratory has demonstrated that diarrheagenic *E. coli* isolates express Mat/ECP fimbriae also at 37°C in the eukaryotic-cell culture medium DMEM (Avelino *et al.*, 2010; Blackburn *et al.*, 2009; Hernandes *et al.*, 2011; Rendón *et al.*, 2007; Saldaña *et al.*, 2009), and the presence of 5% CO_2 atmosphere was shown to further stimulate the fimbriae expression (Rendón *et al.*, 2007). In order to define those specific environmental cues that could promote Mat fimbriae expression in the meningitis isolate IHE 3034, I evaluated a range of laboratory conditions, which may be relevant to bacterial colonization of warm-blooded hosts. DMEM in the absence and presence of 5% CO_2 had no effect on the *matB* transcription or surface expression of Mat fimbriae at 37°C by IHE 3034 (Fig. 1, 2 and S5 of II). In contrast, IHE 3034 bacteria grown at 37°C in LB at low pH (pH 5.5) or in LB containing elevated levels of acetate (50-100 mM) demonstrated increased *mat* promoter activity, *matB* transcription and *matB* mRNA half-life, and the production of the Mat fimbriae (Fig. 1, 2, 9 and S5 of II). The presence of Mat fimbrial filaments on the bacteria after growth at acidic pH (pH 5.5) at 37°C is shown in Figure 9.

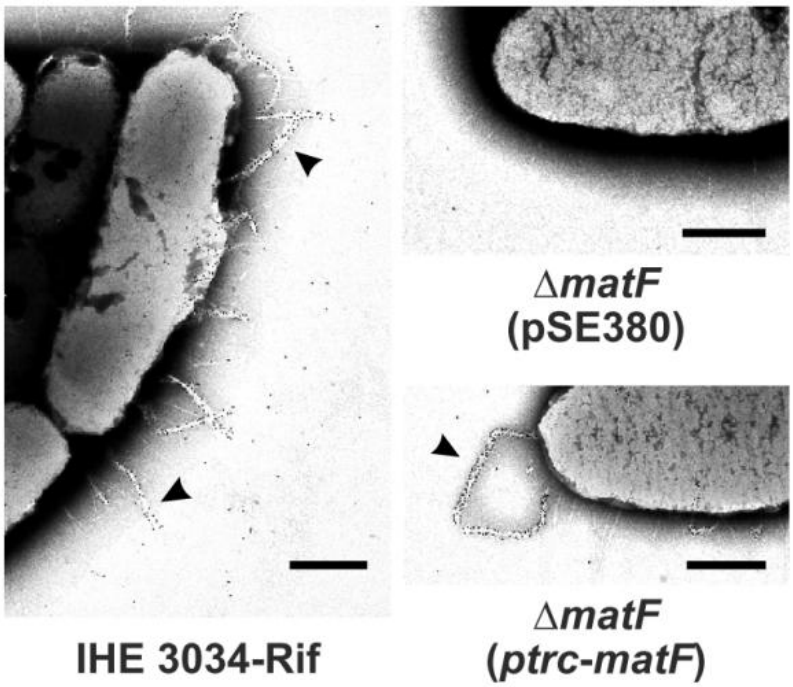

Figure 9. Immunoelectron microscopy of Mat fimbriae. The strain IHE 3034-Rif and the Mat fimbriae-deficient *matF* deletion derivative harbouring pSE380 vector or IPTG-inducible *matF* expression plasmid were cultivated in LB with 0.1 M MES (pH 5.5) at 37°C, and the surface-exposed Mat fimbriae were labeled using anti-Mat antibodies and protein A conjugated to 10-nm gold particles. Arrowheads point to representative gold-labeled Mat fimbriae. Size bars, 0.5 µm.

Intestinal and vaginal bifidobacteria excrete large concentrations of acetate and other SCFAs as the major end products of carbohydrate metabolism, and I showed by serum agglutination that a conditioned medium of *Bifidobacterium longum* cells cultured in acetate-depleted MRS broth promotes Mat expression in IHE 3034 at 37°C. This indicates that *E. coli* is able to express Mat fimbriae at the anatomical sites, the birth canal and intestine. However, the

host-associated cues, acidic pH and elevated levels of acetate or conditioned medium of *B. longum* did not promote biofilm formation of IHE 3034 at 37°C (data not shown).

Further, I tested by whole-cell ELISA analysis whether other strains of phylogenetic group B2 than IHE 3034 are able to induce Mat fimbriae expression under these two host-associated signals and/or low growth temperature (Fig. 2 of II). At 20°C in LB, each B2 strain was able to express the fimbriae, whereas at 37°C the B2 strains showed a considerable degree of variation in responses to acidic pH and high acetate concentration. More than half of the B2 strains (7/11) expressed the fimbriae when cultured at pH 5.5, and several (5/11) also responded to the addition of 0.1 M acetate. In contrast, two tested B1 strains and the A phylogroup strain MG1655 were negative for Mat fimbriation under the examined conditions. These results indicate that environmentally-dependent expression of the adhesive Mat fimbriae involves phylogenetic group -associated differences in regulation. Our results suggest that capability of Mat expression is a common property of B2 isolates. According to current classification, diarrheagenic *E. coli* isolates are found in various phylogenetic groups other than B2 (Escobar-Páramo *et al.*, 2004), whereas extraintestinal isolates cluster mainly into the B2 group (Moulin-Schouleur *et al.*, 2007; Picard *et al.*, 1999). Given that EHEC strain EDL933 express the fimbriae after growth in DMEM at either 26°C or 37°C, but not in LB (Rendón *et al.*, 2007), the Mat fimbriation status seems to rely on different cues in diarrheagenic and extraintestinal *E. coli* strains and may reflect adaptation to specific habitats.

4.2.8 Differential expression of the *mat* operon is due to polymorphisms in the *mat* regulatory region (II)

An intact *mat* gene cluster is present at an identical chromosomal location in 41 out of 45 (91%) completely sequenced *E. coli* genomes representing different phylogenetic groups and pathovars. In an attempt to identify molecular signatures associated with the differential expression of Mat fimbriae in subsets of *E. coli*, I performed a comparative nucleotide sequence analysis on 48 *mat* upstream regulatory regions, including those from publicly available and seven in-house sequences from the strains used in the whole-cell ELISA assay (see Chapter 4.2.7; and Table 7). The analysis revealed two discrete promoter lineages that associate with different phylogenetic groups of *E. coli* (Fig. 2 of II). One promoter lineage contained the isolates of the phylogenetic group B2, D, and the E groups; and the other the A and B1 strains. In contrast, the coding regions of *matA-F* genes showed phylogroup-independent sequence conservation and low degree of strain-specific polymorphisms (Fig. S1 and Table S1 of II). Low ratios of non-synonymous (dN, amino acid replacement) to synonymous (dS) silent changes in the *mat* genes showed characteristics of housekeeping genes (dN/dS<<1; Boyd & Hartl, 1998; Gomez-Duarte *et al.*, 2007), suggesting that *matA-F* genes are evolving under strong purifying selection and purged to preserve their structure and function.

IHE 3034 and MG1655 represent extreme cases according to the Mat surface expression and *mat* sequence analyses. In addition, multi-copy *lacZ* reporter analysis in IHE 3034-Rif background showed that the β-galactosidase activity directed by the pA_{MG1655}-*lacZ* fusion, that encompasses 608 bp of the *matA* upstream region and the first 81bp of *matA* derived

from MG1655, was only 20% of the activity obtained with the $pA_{\text{IHE 3034}}$ promoter fusion (Fig. 5 of II). This result supports the idea that the phylogroup-specific and environmentally-dependent expression of *mat* genes could be assigned to the lineage-restricted polymorphisms in the *mat* regulatory region. To ascertain whether the the promoter sequence variation is a primary determinant of the differential *mat* expression, and thus the reason for the crypticity of the *mat* gene cluster in MG1655, I constructed a series of promoter swapping strains where *mat* regulatory regions were replaced between the chromosomes of IHE 3034 and MG1655 (Fig. 5 of II). Complete loss of fimbriation in IHE 3034 by changing the proximal ~350 base pairs of the upstream regulatory region, that contains the transcriptional start sites T1 and T2 in IHE 3034 and is almost completely conserved in isolates of the B2/D/E lineage, into MG1655-based sequences confirmed that the mat_{MG1655} promoter contains defects that disrupt the fimbriae expression. However, the replacement of this defective promoter region of MG1655 to the corresponding IHE 3034 sequence did not support Mat fimbriation in MG1655 background. In contrast, swapping of an extended 0.6 kb upstream regulatory region containing all three transcription start points enabled crossing of the repression exerted by H-NS and introduced Mat fimbriation capacity into MG1655. This gain-of-function mutagenesis demonstrates that all *mat* genes are functional in MG1655 and the strain has retained the core function of MatA to control *mat* expression. Taken together, our results reveal that the genetic basis of temperature-dependent and phylogenetic group –associated Mat fimbriation phenotype results mainly from *cis*-variations influencing particular promoter architecture. On the other hand, Mat expression response of B2 strains to acidic pH and elevated levels of acetate appears to require strain-specific transcription factors and/or additional regulatory elements beyond 0.6 kb *mat* promoter and *matA*.

5 Conclusions

This study addressed the functional characterization of the common colonization factor Mat fimbriae of *E. coli* and the molecular mechanisms of the regulation of these surface appendages. The results of this thesis work give a novel example of intra-species divergent evolution of chromosomal regulatory DNA element that controls a common but differentially expressed adhesin in a bacterial species associated with various ecological niches or pathogenetic outcomes. We showed that the expression of the *mat* gene cluster of *E. coli* is favored under defined *in vitro* conditions exclusively by isolates of the virulence-associated phylogenetic group B2, including the meningitis strain IHE 3034, but not by the phylogroup B1 and A strains, in a manner dependent on decreased temperature or the host-niche cues acidic pH and elevated levels of acetate. The heterogeneity in Mat fimbriae expression likely offers selective advantage to the bacterium, and might reflect adaptation to specific habitats by subpopulations. Our results demonstrate that the differential expression of the *mat* operon is largely due to polymorphisms in the *mat* upstream regulatory region that falls into two main lineages associated with phylogenetic grouping of *E. coli*, whereas the coding regions of the operon are highly conserved in the genomes. We suggest that the six-gene *mat* operon is an early acquisition in the evolution of *E. coli* and maintained across *E. coli* isolates to facilitate colonization, transmission, and survival at ecological niches that require differential gene expression.

In this thesis work, Mat fimbria is defined as a novel determinant for temperature-dependent biofilm formation in *E. coli*. The protective mode of growth as a biofilm represents a fundamental bacterial survival strategy in changing and severe conditions (Hall-Stoodley *et al.*, 2004). The primary habitat of *E. coli* is in the intestinal tract of warm-blooded animals, but the bacterium also survives and proliferates in water, sediment, and soil (Ishii *et al.*, 2006; Power *et al.*, 2005; Savageau M.A., 1983; Walk *et al.*, 2007). The ability of *E. coli* to exploit the Mat fimbria as an efficient biofilm-promoting factor on abiotic surfaces at low temperatures is an important phenotypic trait for adaptation and persistence outside the warm-blooded host. Moreover, Avelino *et al.* (2010) showed the presence of Mat/ECP fimbriae on a biofilm-growing enteroaggregative *E. coli* (EAEC) 042 cells cultivated in DMEM at 37°C. Thus, biofilm formation facilitated by Mat fimbriae may also play a role in the host mucosal habitats of *E. coli*, namely the intestine and birth canal, that are dominated by multispecies biofilm communities (Macfarlane & Macfarlane, 2006; Macfarlane *et al.*, 2011; Palestrant *et al.*, 2004; Swidsinski *et al.*, 2005) and contains stimulatory signals (low pH, high acetate concentration) for *mat* gene expression. Enhancement of Mat fimbriation as a result of acetate produced by *B. longum* is an example of how bacteria of the human microbiota affect surface structure of *E. coli*.

The present study demonstrates that the transcriptional and post-transcriptional regulation of *mat* gene expression is complex, with overlapping regulatory circuits. Using extensive genetic studies, three regulators, MatA, RcsB, and H-NS, as well as endoribonuclease RNase E that participate in this process were identified. Based on the results of this thesis work, I propose a model for the *mat* regulatory network (Fig. 10) that allows the bacterium to oscillate between a motile, planktonic existence and an adherent, sessile state. In this model, the production of host-cell-adhesive and biofilm-promoting Mat fimbriae is promoted under

conditions of low temperature, acidity, or presence of acetate. MatA and RcsB act as activators that induce transcription initiation from the *mat* promoter, and their activity results in transcription of a polycistronic message covering *matABCDEF* genes. A MatA-mediated positive feedback loop is responsible for counteraction of H-NS –mediated repression of the *mat* locus. In addition, MatA enhances the stability of the processed *matB* mRNA, thus participating in the post-transcriptional regulation of *mat* expression. A second point of post-transcriptional control is provided by the RNase E that initiates the degradation of *matB* transcript at nonpermissive conditions. To avoid simultaneous surface expression of adhesive Mat fimbriae and motility-enhancing flagella, MatA together with RcsB and H-NS forms a regulatory network that conversely coordinates these traits. When the bacterium initiates Mat fimbriae expression, MatA and RcsB bind to the promoter region of flagellar master operon *flhDC*, and consequently prevent flagellum synthesis, and thereby also motility and taxis.

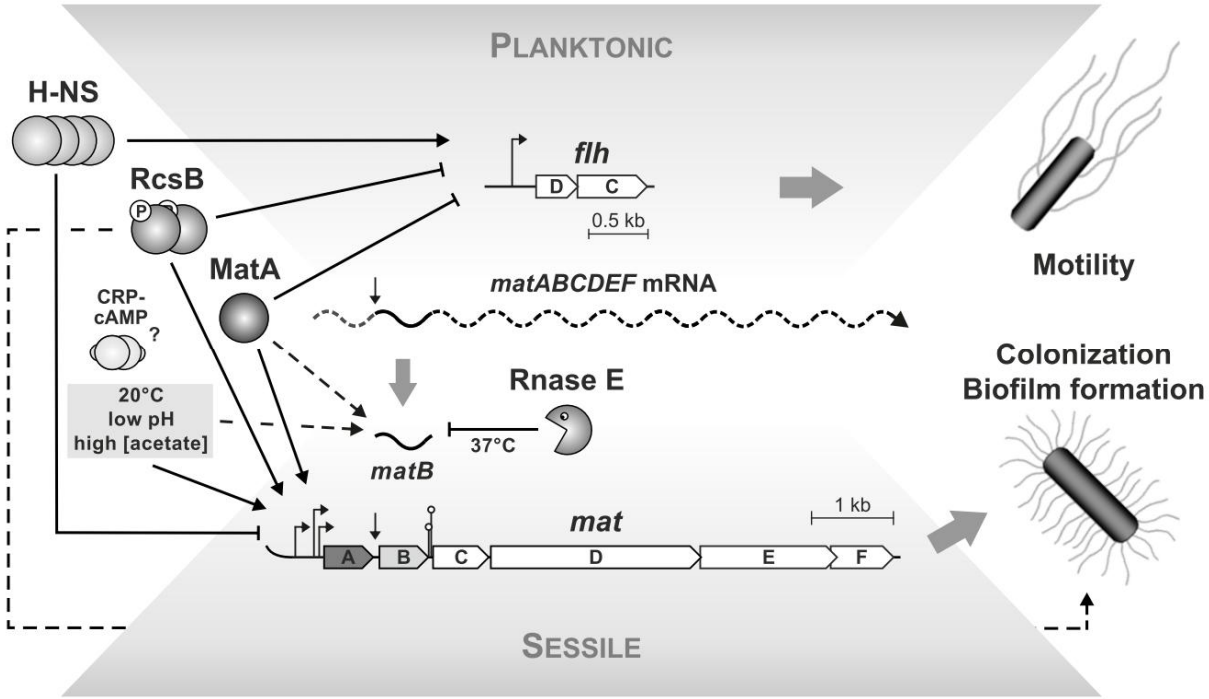

Figure 10. Model depicting the regulation of the *mat* expression that facilitates a motile-to-adhesive transition in *E. coli*. The *mat* gene cluster and flagellar master operon *flhDC* are linked via two regulatory feedforward loop circuits, RcsB-MatA-*flhDC* and H-NS-MatA-*flhDC*. Thus, the fimbriae activator MatA exerts a dual regulatory function on the choice of two opposite functions, Mat fimbriae-mediated adherence and flagella-driven motility. An unstable polycistronic transcript, *matABCDEF*, is transcribed from the *mat* operon. The fimbrillin *matB* mRNA is processed from the larger precursor mRNA and is preferentially stabilized at low temperature, and after a temperature shift to 37°C the transcript becomes susceptible to RNase E-mediated degradation. In meningitis isolate IHE 3034, low temperature, acidic pH, and elevated levels of acetate induce Mat fimbria expression by activating the *mat* promoter and by stabilizing the fimbrillin mRNA. In addition to MatA, RcsB, and H-NS, CRP-cAMP is able to bind to the *mat* regulatory DNA. The role of CRP-cAMP binding, however, remains open at the present. Right-angled arrows indicate transcription start sites, vertical black arrows show the 5' end of monocistronic *matB* mRNA, and stem-loop symbols indicate the putative transcriptional terminators downstream of *matB*. Arrowheads indicate activation, and blunt ends indicate repression. Direct regulation is shown as continuous lines, and indirect effects are represented as dotted lines.

The regulatory model in Figure 10 provides the primary mechanisms underlying the expression of the *mat* operon. Undoubtedly, our catalogue of *trans*-acting factors is incomplete, and further studies are required to gain a coherent picture of regulatory network of *mat* system. For example, a global transcription factor Lrp has a leusine-sensitive binding site in the *mat* promoter, as discovered by ChIP-chip analysis by Cho *et al.* (2008), and subsequent genome-wide gene expression analysis indicated that Lrp activates the transcription of *matA* in the absence of exogenous leucine. Moreover, the transcription factors involved in the *mat* promoter responses to low pH and high acetate concentration remains to be identified. The elucidation of adhesin protein(s) in the Mat filament, the target receptor(s) in the host tissues as well as the *in vivo* regulation of Mat expression are of the utmost importance in order to better understand the function of the Mat fimbriae. Another interesting subject is the disruption or complete loss of *mat* operon in pathogenic *Shigella*, a facultative intracellular pathogen, and in a small subset of *E. coli*. It is tempting to speculate that these inactivations represent genetic trade-offs caused by antagonistic pleiotrophy. The evolutionary model of antagonistic pleiotropy predicts that traits beneficial for one niche are detrimental in another environment (Cooper & Lenski, 2000; Rose & Charlesworth, 1980). Thus, in the case of *Shigella* that invades the colonic epithelium, inactivation of the *mat* genes may be an example of pathoadaptive mutation, as seen for the lysine decarboxylase gene *cadA* (Day *et al.*, 2001; Maurelli *et al.*, 1998) and curli loci (Sakellaris *et al.*, 2000), that results in enhanced pathogenicity.

Taken together, our findings on the fine-tuned mechanisms of differential *mat* expression are important in demonstrating genomic alterations in regulatory DNA that affect regulation of conserved and ancestral *E. coli* genes important for host interactions and for living in a biofilm, and thus exemplify the evolution of phenotypic diversity within bacterial species. We are only beginning to comprehend that evolutionary adaptation to pathogenic lifestyles involves, in addition to gain and loss of genes, regulatory variations in expression of genes present both in pathogens and commensals. The data provided herein will drive the research on regulation of *E. coli* virulence and fitness factors towards a more complete understanding of the evolution of bacterial virulence and of the concept of pathoadaptive variation of genes and regulatory DNA. Considering the prevalence of *mat* locus in *E. coli* genomes, the adhesive-properties of Mat filaments, and the regulatory connection to the flagella system, we suggest that Mat-mediated cell-adhesion and biofilm formation advance the evolutionary success of *E. coli*, especially the extraintestinal isolates of the virulence-associated B2 group, by increasing the colonization and transmission capacity as well as the persistence in the environment.

6 Acknowledgements

This thesis work was carried out at the Division of General Microbiology, Department of Biosciences at the University of Helsinki. The financial support for the work was provided by Viikki Doctoral Programme in Molecular Biosciences (VGSB), the European Commission under FP5 (Colirisk project), the European Union Network of Excellence (EuroPathoGenomics program), and the Academy of Finland.

First of all, I owe my deepest gratitude to my excellent supervisors, Professor Timo K. Korhonen and Docent Benita Westerlund-Wikström. Thanks to Timo, I got a contract with the ”not-so-straightforward” Mat project and an introduction to the world of fimbriae during my undergraduate studies. Since then, Timo has been a calm and visionary supervisor, always encouraging me, in his own way. I appreciate your keen intelligence and your confidence in me and the project. Benita’s realistic and down-to-earth – and always optimistic – approach has been priceless for me during the years. I thank your assistance in overseeing the project on a daily basis and keeping the office door always open for me in the need of guidance. You have invariably supported and motivated me, pushing me forward. And the constant supply of fresh coffee you provided should not be underestimated either. I would like to thank Professor Kielo Haahtela, the head of the department, for providing the excellent research facilities and organizing the division’s Friday (nowadays Thursday) coffee.

I am deeply grateful to my collaborators, Professor Ulrich Dobrindt and Philippe Bauchart for their valuable contribution to the project and advice during the years. Members of my thesis advisory committee, Professor Jaana Bamford and Professor Harri Savilahti, are acknowledged for their interest and positive attitude as well as the encouraging feedback during our meetings. I would also like to thank Harri and Docent Suvi Taira for the careful, flexible and fast review of my thesis and for valuable and positive comments.

My sincerest thanks also go to all the past and present members of the Adheesio-lab, including Ritva (my mentor of microscopy), Katri (always full of a positive attitude!), Veera (great coffee company), Riikka, Lena, Maiju, Jenni, Päivi, Kaarina, Sanna, and Johanna He. My special thanks go to Maini for valuable brainstorming sessions that taught me the basics of research mentality at the very beginning of my journey in science; and Johanna Ha. for supporting me for my ups and downs in science and for standing the nearly constant buzz of conversation in our shared office. I am also very grateful to the invaluable laboratory staff, especially Raili, Jaana, Merja and Ari, for keeping the laboratory and machines running smoothly and maintaining all the supplies.

Last, I would like to thank all my dear friends and family, including parents-in-law, who have supported me in varying ways outside the lab and academic world. A special mention goes to the Sauna gang “Ready for One!”, Mika, Timo K., and Martti. This assembly has literally helped take my mind off science from time to time. Finally and foremost, I wish to thank my wife, Nina, and our daughter, Milla, for all the love. And for the incredible amount of patience Nina had with me in the past twelve months (or twelve years!). I really am amazed by the confidence that she brings to my life. Nina, you are a constant source of

inspiration in every aspect of my life. You are my muse, my rock - the light and fire of my life.

Helsinki, May 2012
Timo

7 References

Åberg V. & Almqvist F. (2007). Pilicides-small molecules targeting bacterial virulence. *Org Biomol Chem* **5**, 1827-1834.

Abraham J. M., Freitag C. S., Clements J. R., Eisenstein B. I. (1985). An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc Natl Acad Sci U S A* **82**, 5724-5727.

Achtman M., Mercer A., Kusecek B., Pohl A., Heuzenroeder M., Aaronson W., Sutton A., Silver R. P. (1983). Six widespread bacterial clones among *Escherichia coli* K1 isolates. *Infect Immun* **39**, 315-335.

Adlerberth I., Hanson L. A., Svanborg C., Svennerholm A. M., Nordgren S., Wold A. E. (1995). Adhesins of *Escherichia coli* associated with extra-intestinal pathogenicity confer binding to colonic epithelial cells. *Microb Pathog* **18**, 373-385.

Ahmed N., Dobrindt U., Hacker J., Hasnain S. E. (2008). Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. *Nat Rev Microbiol* **6**, 387-394.

Akerley B. J., Cotter P. A., Miller J. F. (1995). Ectopic expression of the flagellar regulon alters development of the Bordetella-host interaction. *Cell* **80**, 611-620.

Allen P. M., Roberts I., Boulnois G. J., Saunders J. R., Hart C. A. (1987). Contribution of capsular polysaccharide and surface properties to virulence of *Escherichia coli* K1. *Infect Immun* **55**, 2662-2668.

Amini S., Goodarzi H., Tavazoie S. (2009). Genetic dissection of an exogenously induced biofilm in laboratory and clinical isolates of *E. coli*. *PLoS Pathog* **5**, e1000432.

Anantha R. P., McVeigh A. L., Lee L. H., Agnew M. K., Cassels F. J., Scott D. A., Whittam T. S., Savarino S. J. (2004). Evolutionary and functional relationships of colonization factor antigen i and other class 5 adhesive fimbriae of enterotoxigenic *Escherichia coli*. *Infect Immun* **72**, 7190-7201.

Anderson G. G., Palermo J. J., Schilling J. D., Roth R., Heuser J., Hultgren S. J. (2003). Intracellular bacterial biofilm-like pods in urinary tract infections. *Science* **301**, 105-107.

Andersson M., Uhlin B. E., Fällman E. (2007). The biomechanical properties of *E. coli* pili for urinary tract attachment reflect the host environment. *Biophys J* **93**, 3008-3014.

Andersson M., Fällman E., Uhlin B. E., Axner O. (2006). A sticky chain model of the elongation and unfolding of *Escherichia coli* P pili under stress. *Biophys J* **90**, 1521-1534.

Andersson M., Axner O., Almqvist F., Uhlin B. E., Fällman E. (2008). Physical properties of biopolymers assessed by optical tweezers: analysis of folding and refolding of bacterial pili. *Chemphyschem* **9**, 221-235.

Archer C. T., Kim J. F., Jeong H., Park J. H., Vickers C. E., Lee S. Y., Nielsen L. K. (2011). The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. *BMC Genomics* **12**, 9.

Arraiano C. M., Andrade J. M., Domingues S., Guinote I. B., Malecki M., Matos R. G., Moreira R. N., Pobre V., Reis F. P.& other authors. (2010). The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol Rev* **34**, 883-923.

Artis D. (2008). Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. *Nat Rev Immunol* **8**, 411-420.

Atlung T. & Ingmer H. (1997). H-NS: a modulator of environmentally regulated gene expression. *Mol Microbiol* **24**, 7-17.

Avelino F., Saldaña Z., Islam S., Monteiro-Neto V., Dall'Agnol M., Eslava C. A., Girón J. A. (2010). The majority of enteroaggregative *Escherichia coli* strains produce the *E. coli* common pilus when adhering to cultured epithelial cells. *Int J Med Microbiol* **300**, 440-448.

Axner O., Andersson M., Björnham O., Castelain M., Klinth J., Koutris E., Schedin S. (2011). Assessing bacterial adhesion on an individual adhesin and single pili level using optical tweezers. *Adv Exp Med Biol* **715**, 301-313.

Babai R., Blum-Oehler G., Stern B. E., Hacker J., Ron E. Z. (1997). Virulence patterns from septicemic *Escherichia coli* O78 strains. *FEMS Microbiol Lett* **149**, 99-105.

Båga M., Göransson M., Normark S., Uhlin B. E. (1988). Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression. *Cell* **52**, 197-206.

Båga M., Göransson M., Normark S., Uhlin B. E. (1985). Transcriptional activation of a Pap pilus virulence operon from uropathogenic *Escherichia coli*. *EMBO J* **4**, 3887-3893.

Bakker D., Willemsen P. T., Simons L. H., van Zijderveld F. G., de Graaf F. K. (1992). Characterization of the antigenic and adhesive properties of FaeG, the major subunit of K88 fimbriae. *Mol Microbiol* **6**, 247-255.

Balsalobre C., Morschhäuser J., Jass J., Hacker J., Uhlin B. E. (2003). Transcriptional analysis of the *sfa* determinant revealing mmRNA processing events in the biogenesis of S fimbriae in pathogenic *Escherichia coli*. *J Bacteriol* **185**, 620-629.

Baorto D. M., Gao Z., Malaviya R., Dustin M. L., van der Merwe A., Lublin D. M., Abraham S. N. (1997). Survival of FimH-expressing enterobacteria in macrophages relies on glycolipid traffic. *Nature* **389**, 636-639.

Barlow D. J. & Thornton J. M. (1988). Helix geometry in proteins. *J Mol Biol* **201**, 601-619.

Bauchart P., Germon P., Brée A., Oswald E., Hacker J., Dobrindt U. (2010). Pathogenomic comparison of human extraintestinal and avian pathogenic *Escherichia coli*—search for factors involved in host specificity or zoonotic potential. *Microb Pathog* **49**, 105-115.

Belasco J. G. (2010). All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat Rev Mol Cell Biol* **11**, 467-478.

Beloin C., Michaelis K., Lindner K., Landini P., Hacker J., Ghigo J. M., Dobrindt U. (2006). The transcriptional antiterminator RfaH represses biofilm formation in *Escherichia coli*. *J Bacteriol* **188**, 1316-1331.

Ben Nasr A., Olsén A., Sjöbring U., Müller-Esterl W., Björck L. (1996). Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing *Escherichia coli*. *Mol Microbiol* **20**, 927-935.

Bendtsen J. D., Nielsen H., von Heijne G., Brunak S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**, 783-795.

Berg O. G. & von Hippel P. H. (1988). Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. *J Mol Biol* **200**, 709-723.

Berger C. N., Billker O., Meyer T. F., Servin A. L., Kansau I. (2004). Differential recognition of members of the carcinoembryonic antigen family by Afa/Dr adhesins of diffusely adhering *Escherichia coli* (Afa/Dr DAEC). *Mol Microbiol* **52**, 963-983.

Berger H., Hacker J., Juarez A., Hughes C., Goebel W. (1982). Cloning of the chromosomal determinants encoding hemolysin production and mannose-resistant hemagglutination in *Escherichia coli*. *J Bacteriol* **152**, 1241-1247.

Bergthorsson U. & Ochman H. (1998). Distribution of chromosome length variation in natural isolates of *Escherichia coli*. *Mol Biol Evol* **15**, 6-16.

Bertin P., Terao E., Lee E. H., Lejeune P., Colson C., Danchin A., Collatz E. (1994). The H-NS protein is involved in the biogenesis of flagella in *Escherichia coli*. *J Bacteriol* **176**, 5537-5540.

Bettelheim K. A. & Lennox-King S. M. (1976). The acquisition of *Escherichia coli* by new-born babies. *Infection* **4**, 174-179.

Bettelheim K. A., Teoh-Chan C. H., Chandler M. E., O'Farrell S. M., Rahamim L., Shaw E. J., Shooter R. A. (1974). Spread of *Escherichia coli* colonizing newborn babies and their mothers. *J Hyg (Lond)* **73**, 383-387.

Bilge S. S., Apostol J. M.,Jr, Aldape M. A., Moseley S. L. (1993). mRNA processing independent of RNase III and RNase E in the expression of the F1845 fimbrial adhesin of *Escherichia coli*. *Proc Natl Acad Sci U S A* **90**, 1455-1459.

Bingen E., Bonacorsi S., Brahimi N., Denamur E., Elion J. (1997). Virulence patterns of *Escherichia coli* K1 strains associated with neonatal meningitis. *J Clin Microbiol* **35**, 2981-2982.

Bingen E., Picard B., Brahimi N., Mathy S., Desjardins P., Elion J., Denamur E. (1998). Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *J Infect Dis* **177**, 642-650.

Blackburn D., Husband A., Saldaña Z., Nada R. A., Klena J., Qadri F., Girón J. A. (2009). Distribution of the *Escherichia coli* common pilus among diverse strains of human enterotoxigenic *E. coli*. *J Clin Microbiol* **47**, 1781-1784.

Blattner F. R., Plunkett G.,3rd, Bloch C. A., Perna N. T., Burland V., Riley M., Collado-Vides J., Glasner J. D., Rode C. K.& other authors. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1462.

Bloch C. A., Stocker B. A., Orndorff P. E. (1992). A key role for type 1 pili in enterobacterial communicability. *Mol Microbiol* **6**, 697-701.

Blomfield I. C., Calie P. J., Eberhardt K. J., McClain M. S., Eisenstein B. I. (1993). Lrp stimulates phase variation of type 1 fimbriation in *Escherichia coli* K-12. *J Bacteriol* **175**, 27-36.

Blum E., Carpousis A. J., Higgins C. F. (1999). Polyadenylation promotes degradation of 3'-structured RNA by the *Escherichia coli* mRNA degradosome *in vitro*. *J Biol Chem* **274**, 4009-4016.

Blyn L. B., Braaten B. A., Low D. A. (1990). Regulation of *pap* pilin phase variation by a mechanism involving differential dam methylation states. *EMBO J* **9**, 4045-4054.

Bolivar F., Rodriguez R. L., Greene P. J., Betlach M. C., Heyneker H. L., Boyer H. W., Crosa J. H., Falkow S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**, 95-113.

Bollinger R. R., Everett M. L., Wahl S. D., Lee Y. H., Orndorff P. E., Parker W. (2006). Secretory IgA and mucin-mediated biofilm formation by environmental strains of *Escherichia coli*: role of type 1 pili. *Mol Immunol* **43**, 378-387.

Bonacorsi S. & Bingen E. (2005). Molecular epidemiology of *Escherichia coli* causing neonatal meningitis. *Int J Med Microbiol* **295**, 373-381.

Bortolussi R., Ferrieri P., Quie P. G. (1983). Influence of growth temperature of *Escherichia coli* on K1 capsular antigen production and resistance to opsonization. *Infect Immun* **39**, 1136-1141.

Boudeau J., Glasser A. L., Masseret E., Joly B., Darfeuille-Michaud A. (1999). Invasive ability of an *Escherichia coli* strain isolated from the ileal mucosa of a patient with Crohn's disease. *Infect Immun* **67**, 4499-4509.

Bouffartigues E., Buckle M., Badaut C., Travers A., Rimsky S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat Struct Mol Biol* **14**, 441-448.

Bouvier M., Sharma C. M., Mika F., Nierhaus K. H., Vogel J. (2008). Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* **32**, 827-837.

Boyd E. F. & Hartl D. L. (1998). Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. *J Bacteriol* **180**, 1159-1165.

Bray J. (1945). Isolation of antigenically homogeneous strains of *Bact. coli neapolitanum* from summer diarrhœa of infants. *J Pathol Bacteriol* **57**, 239-247.

Bricker A. L. & Belasco J. G. (1999). Importance of a 5' stem-loop for longevity of *papA* mRNA in *Escherichia coli*. *J Bacteriol* **181**, 3587-3590.

Brill J. A., Quinlan-Walshe C., Gottesman S. (1988). Fine-structure mapping and identification of two regulators of capsule synthesis in *Escherichia coli* K-12. *J Bacteriol* **170**, 2599-2611.

Brinton C. C.,Jr. (1959). Non-flagellar appendages of bacteria. *Nature* **183**, 782-786.

Brown J., Faulds-Pain A., Aldridge P. (2009). The Coordination of Flagellar Gene Expression and the Flagellar Assembly Pathway. In *Pili and Flagella: Current Research and Future Trends,* pp. 99-120. Edited by K. Jarrell. Norfolk, UK: Caister Academic Press.

Brzuszkiewicz E., Thürmer A., Schuldes J., Leimbach A., Liesegang H., Meyer F. D., Boelter J., Petersen H., Gottschalk G., Daniel R. (2011). Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Arch Microbiol* **193**, 883-891.

Buchberger A., Schröder H., Hesterkamp T., Schönfeld H. J., Bukau B. (1996). Substrate shuttling between the DnaK and GroEL systems indicates a chaperone network promoting protein folding. *J Mol Biol* **261**, 328-333.

Bühler T., Hoschützky H., Jann K. (1991). Analysis of colonization factor antigen I, an adhesin of enterotoxigenic *Escherichia coli* O78:H11: fimbrial morphology and location of the receptor-binding site. *Infect Immun* **59**, 3876-3882.

Bury-Moné S., Nomane Y., Reymond N., Barbet R., Jacquet E., Imbeaud S., Jacq A., Bouloc P. (2009). Global analysis of extracytoplasmic stress signaling in *Escherichia coli*. *PLoS Genet* **5**, e1000651.

Callaghan A. J., Marcaida M. J., Stead J. A., McDowall K. J., Scott W. G., Luisi B. F. (2005). Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* **437**, 1187-1191.

Carpousis A. J. (2007). The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* **61**, 71-87.

Casadesús J. & Low D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* **70**, 830-856.

Cegelski L., Pinkner J. S., Hammer N. D., Cusumano C. K., Hung C. S., Chorell E., Åberg V., Walker J. N., Seed P. C.& other authors. (2009). Small-molecule inhibitors target *Escherichia coli* amyloid biogenesis and biofilm formation. *Nat Chem Biol* **5**, 913-919.

Chaudhuri R. R., Sebaihia M., Hobman J. L., Webber M. A., Leyton D. L., Goldberg M. D., Cunningham A. F., Scott-Tucker A., Ferguson P. R.& other authors. (2010). Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. *PLoS One* **5**, e8801.

Chen S. L., Hung C. S., Pinkner J. S., Walker J. N., Cusumano C. K., Li Z., Bouckaert J., Gordon J. I., Hultgren S. J. (2009). Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding. *Proc Natl Acad Sci U S A* **106**, 22439-22444.

Cherepanov P. P. & Wackernagel W. (1995). Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9-14.

Cho B. K., Barrett C. L., Knight E. M., Park Y. S., Palsson B. Ø. (2008). Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc Natl Acad Sci U S A* **105**, 19462-19467.

Clarke D. J. (2010). The Rcs phosphorelay: more than just a two-component pathway. *Future Microbiol* **5**, 1173-1184.

Clavijo A. P., Bai J., Gómez-Duarte O. G. (2010). The Longus type IV pilus of enterotoxigenic *Escherichia coli* (ETEC) mediates bacterial self-aggregation and protection from antimicrobial agents. *Microb Pathog* **48**, 230-238.

Cleary J., Lai L. C., Shaw R. K., Straatman-Iwanowska A., Donnenberg M. S., Frankel G., Knutton S. (2004). Enteropathogenic *Escherichia coli* (EPEC) adhesion to intestinal epithelial cells: role of bundle-forming pili (BFP), EspA filaments and intimin. *Microbiology* **150**, 527-538.

Clegg S. & Hughes K. T. (2002). FimZ is a molecular link between sticking and swimming in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **184**, 1209-1213.

Clegg S., Wilson J., Johnson J. (2011). More than one way to control hair growth: regulatory mechanisms in enterobacteria that affect fimbriae assembled by the chaperone/usher pathway. *J Bacteriol* **193**, 2081-2088.

Clermont O., Bonacorsi S., Bingen E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* **66**, 4555-4558.

Clermont O., Lescat M., O'Brien C. L., Gordon D. M., Tenaillon O., Denamur E. (2008). Evidence for a human-specific *Escherichia coli* clone. *Environ Microbiol* **10**, 1000-1006.

Connell I., Agace W., Klemm P., Schembri M., Mårild S., Svanborg C. (1996). Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proc Natl Acad Sci U S A* **93**, 9827-9832.

Cooper V. S. & Lenski R. E. (2000). The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* **407**, 736-739.

Corcoran C. P. & Dorman C. J. (2009). DNA relaxation-dependent phase biasing of the *fim* genetic switch in *Escherichia coli* depends on the interplay of H-NS, IHF and LRP. *Mol Microbiol* **74**, 1071-1082.

Costello E. K., Lauber C. L., Hamady M., Fierer N., Gordon J. I., Knight R. (2009). Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694-1697.

Cross A. S., Kim K. S., Wright D. C., Sadoff J. C., Gemski P. (1986). Role of lipopolysaccharide and capsule in the serum resistance of bacteremic strains of *Escherichia coli*. *J Infect Dis* **154**, 497-503.

Croxen M. A. & Finlay B. B. (2010). Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol* **8**, 26-38.

Czeczulin J. R., Balepur S., Hicks S., Phillips A., Hall R., Kothary M. H., Navarro-Garcia F., Nataro J. P. (1997). Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infect Immun* **65**, 4135-4145.

Daigle F., Forget C., Martin C., Drolet M., Tessier M. C., Dezfulian H., Harel J. (2000). Effects of global regulatory proteins and environmental conditions on fimbrial gene expression of $F165_1$ and $F165_2$ produced by *Escherichia coli* causing septicaemia in pigs. *Res Microbiol* **151**, 563-574.

Danese P. N., Pratt L. A., Kolter R. (2000). Exopolysaccharide production is required for development of *Escherichia coli* K-12 biofilm architecture. *J Bacteriol* **182**, 3593-3596.

Daniel A., Singh A., Crowther L. J., Fernandes P. J., Schreiber W., Donnenberg M. S. (2006). Interaction and localization studies of enteropathogenic *Escherichia coli* type IV bundle-forming pilus outer membrane components. *Microbiology* **152**, 2405-2420.

Darfeuille F., Unoson C., Vogel J., Wagner E. G. (2007). An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell* **26**, 381-392.

Darfeuille-Michaud A., Neut C., Barnich N., Lederman E., Di Martino P., Desreumaux P., Gambiez L., Joly B., Cortot A., Colombel J. F. (1998). Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology* **115**, 1405-1413.

Das M., Hart-Van Tassell A., Urvil P. T., Lea S., Pettigrew D., Anderson K. L., Samet A., Kur J., Matthews S.& other authors. (2005). Hydrophilic domain II of *Escherichia coli* Dr fimbriae facilitates cell invasion. *Infect Immun* **73**, 6119-6126.

Datsenko K. A. & Wanner B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645.

d'Aubenton Carafa Y., Brody E., Thermes C. (1990). Prediction of rho-independent *Escherichia coli* transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* **216**, 835-858.

Day W. A.,Jr, Fernández R. E., Maurelli A. T. (2001). Pathoadaptive mutations that enhance virulence: genetic organization of the *cadA* regions of *Shigella* spp. *Infect Immun* **69**, 7471-7480.

de Crombrugghe B., Busby S., Buc H. (1984). Cyclic AMP receptor protein: role in transcription activation. *Science* **224**, 831-838.

de Muinck E. J., Øien T., Storrø O., Johnsen R., Stenseth N. C., Rønningen K. S., Rudi K. (2011). Diversity, transmission and persistence of *Escherichia coli* in a cohort of mothers and their infants. *Environ Microbiol Rep* **3**, 352-359.

Deana A. & Belasco J. G. (2005). Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev* **19**, 2526-2533.

Deininger K. N., Horikawa A., Kitko R. D., Tatsumi R., Rosner J. L., Wachi M., Slonczewski J. L. (2011). A requirement of TolC and MDR efflux pumps for acid adaptation and GadAB induction in *Escherichia coli*. *PLoS One* **6**, e18960.

Deutscher M. P. (2006). Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res* **34**, 659-666.

Dho-Moulin M. & Fairbrother J. M. (1999). Avian pathogenic *Escherichia coli* (APEC). *Vet Res* **30**, 299-316.

Dietzman D. E., Fischer G. W., Schoenknecht F. D. (1974). Neonatal *Escherichia coli* septicemia—bacterial counts in blood. *J Pediatr* **85**, 128-130.

Dobrindt U. (2005). (Patho-)Genomics of *Escherichia coli*. *Int J Med Microbiol* **295**, 357-371.

Dobrindt U., Chowdary M. G., Krumbholz G., Hacker J. (2010). Genome dynamics and its impact on evolution of *Escherichia coli*. *Med Microbiol Immunol* **199**, 145-154.

Dobrindt U., Agerer F., Michaelis K., Janka A., Buchrieser C., Samuelson M., Svanborg C., Gottschalk G., Karch H., Hacker J. (2003). Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J Bacteriol* **185**, 1831-1840.

Dodd D. C. & Eisenstein B. I. (1984). Kinetic analysis of the synthesis and assembly of type 1 fimbriae of *Escherichia coli*. *J Bacteriol* **160**, 227-232.

Domka J., Lee J., Bansal T., Wood T. K. (2007). Temporal gene-expression in *Escherichia coli* K-12 biofilms. *Environ Microbiol* **9**, 332-346.

Donnenberg M. S. & Kaper J. B. (1991). Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect Immun* **59**, 4310-4317.

Dorman C. J. (2004). H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* **2**, 391-400.

Dubnau D. & Losick R. (2006). Bistability in bacteria. *Mol Microbiol* **61**, 564-572.

Duguid J. P., Smith I. W., Dempster G., Edmunds P. N. (1955). Non-flagellar filamentous appendages (fimbriæ) and haemagglutinating activity in B*acterium coli* . *J Pathol Bacteriol* **70**, 335-348.

Duncan M. J., Mann E. L., Cohen M. S., Ofek I., Sharon N., Abraham S. N. (2005). The distinct binding specificities exhibited by enterobacterial type 1 fimbriae are determined by their fimbrial shafts. *J Biol Chem* **280**, 37707-37716.

Duriez P., Clermont O., Bonacorsi S., Bingen E., Chaventré A., Elion J., Picard B., Denamur E. (2001). Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* **147**, 1671-1676.

Eckburg P. B., Bik E. M., Bernstein C. N., Purdom E., Dethlefsen L., Sargent M., Gill S. R., Nelson K. E., Relman D. A. (2005). Diversity of the human intestinal microbial flora. *Science* **308**, 1635-1638.

Edén C. S. & Hansson H. A. (1978). *Escherichia coli* pili as possible mediators of attachment to human urinary tract epithelial cells. *Infect Immun* **21**, 229-237.

Edwards R. A. & Puente J. L. (1998). Fimbrial expression in enteric bacteria: a critical step in intestinal pathogenesis. *Trends Microbiol* **6**, 282-287.

Edwards R. A. & Schifferli D. M. (1997). Differential regulation of *fasA* and *fasH* expression of *Escherichia coli* 987P fimbriae by environmental cues. *Mol Microbiol* **25**, 797-809.

Eisenstein B. I., Sweet D. S., Vaughn V., Friedman D. I. (1987). Integration host factor is required for the DNA inversion that controls phase variation in *Escherichia coli*. *Proc Natl Acad Sci U S A* **84**, 6506-6510.

Erdem A. L., Avelino F., Xicohtencatl-Cortes J., Girón J. A. (2007). Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing *Escherichia coli*. *J Bacteriol* **189**, 7426-7435.

Escherich T. (1885). Die darmbakterien des neugeborenen und säuglings. *Fortschr Med* **16**, 515-522; 547–554.

Escobar-Páramo P., Clermont O., Blanc-Potard A. B., Bui H., Le Bouguénec C., Denamur E. (2004). A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol Biol Evol* **21**, 1085-1094.

Escobar-Páramo P., Le Menac'h A., Le Gall T., Amorin C., Gouriou S., Picard B., Skurnik D., Denamur E. (2006). Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environ Microbiol* **8**, 1975-1984.

Evans D. G., Silver R. P., Evans D. J.,Jr, Chase D. G., Gorbach S. L. (1975). Plasmid-controlled colonization factor associated with virulence in *Esherichia coli* enterotoxigenic for humans. *Infect Immun* **12**, 656-667.

Ewers C., Li G., Wilking H., Kieβling S., Alt K., Antáo E. M., Laturnus C., Diehl I., Glodde S.& other authors. (2007). Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: how closely related are they? *Int J Med Microbiol* **297**, 163-176.

Farfan M. J., Cantero L., Vidal R., Botkin D. J., Torres A. G. (2011). Long polar fimbriae of enterohemorrhagic *Escherichia coli* O157:H7 bind to extracellular matrix proteins. *Infect Immun* **79**, 3744-3750.

Farmer J. J.,3rd, Fanning G. R., Davis B. R., O'Hara C. M., Riddle C., Hickman-Brenner F. W., Asbury M. A., Lowery V. A.,3rd, Brenner D. J. (1985). *Escherichia fergusonii* and *Enterobacter taylorae*, two new species of *Enterobacteriaceae* isolated from clinical specimens. *J Clin Microbiol* **21**, 77-81.

Ferrières L. & Clarke D. J. (2003). The RcsC sensor kinase is required for normal biofilm formation in *Escherichia coli* K-12 and controls the expression of a regulon in response to growth on a solid surface. *Mol Microbiol* **50**, 1665-1682.

Ferrieri P., Burke B., Nelson J. (1980). Production of bacteremia and meningitis in infant rats with group B streptococcal serotypes. *Infect Immun* **27**, 1023-1032.

Finne J. (1982). Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. *J Biol Chem* **257**, 11966-11970.

Finne J., Finne U., Deagostini-Bazin H., Goridis C. (1983). Occurrence of α2-8 linked polysialosyl units in a neural cell adhesion molecule. *Biochem Biophys Res Commun* **112**, 482-487.

Flemming H. C. & Wingender J. (2010). The biofilm matrix. *Nat Rev Microbiol* **8**, 623-633.

Forero M., Yakovenko O., Sokurenko E. V., Thomas W. E., Vogel V. (2006). Uncoiling mechanics of *Escherichia coli* type I fimbriae are optimized for catch bonds. *PLoS Biol* **4**, e298.

Francez-Charlot A., Laugel B., Van Gemert A., Dubarry N., Wiorowski F., Castanié-Cornet M. P., Gutierrez C., Cam K. (2003). RcsCDB His-Asp phosphorelay system negatively regulates the *flhDC* operon in *Escherichia coli*. *Mol Microbiol* **49**, 823-832.

Fredericks C. E., Shibata S., Aizawa S., Reimann S. A., Wolfe A. J. (2006). Acetyl phosphate-sensitive regulation of flagellar biogenesis and capsular biosynthesis depends on the Rcs phosphorelay. *Mol Microbiol* **61**, 734-747.

Freitag C. S., Abraham J. M., Clements J. R., Eisenstein B. I. (1985). Genetic analysis of the phase variation control of expression of type 1 fimbriae in *Escherichia coli*. *J Bacteriol* **162**, 668-675.

Freter R. & Abrams G. D. (1972). Function of various intestinal bacteria in converting germfree mice to the normal state. *Infect Immun* **6**, 119-126.

Fronzes R., Remaut H., Waksman G. (2008). Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. *EMBO J* **27**, 2271-2280.

Furyk J. S., Swann O., Molyneux E. (2011). Systematic review: neonatal meningitis in the developing world. *Trop Med Int Health* **16**, 672-679.

Gally D. L., Leathart J., Blomfield I. C. (1996). Interaction of FimB and FimE with the *fim* switch that controls the phase variation of type 1 fimbriae in *Escherichia coli* K-12. *Mol Microbiol* **21**, 725-738.

Gally D. L., Rucker T. J., Blomfield I. C. (1994). The leucine-responsive regulatory protein binds to the *fim* switch to control phase variation of type 1 fimbrial expression in *Escherichia coli* K-12. *J Bacteriol* **176**, 5665-5672.

Gally D. L., Bogan J. A., Eisenstein B. I., Blomfield I. C. (1993). Environmental regulation of the *fim* switch controlling type 1 fimbrial phase variation in *Escherichia coli* K-12: effects of temperature and media. *J Bacteriol* **175**, 6186-6193.

Gao R. & Stock A. M. (2009). Biological insights from structures of two-component proteins. *Annu Rev Microbiol* **63**, 133-154.

Gardel C. L. & Mekalanos J. J. (1996). Alterations in *Vibrio cholerae* motility phenotypes correlate with changes in virulence factor expression. *Infect Immun* **64**, 2246-2255.

Garnett J. A., Martínez-Santos V. I., Saldaña Z., Pape T., Hawthorne W., Chan J., Simpson P. J., Cota E., Puente J. L., Girón J. A., Matthews S. (2012). Structural insights into the biogenesis and biofilm formation by the *Escherichia coli* common pilus. *Proc Natl Acad Sci U S A* **109**, 3950-3955.

Gaschignard J., Levy C., Romain O., Cohen R., Bingen E., Aujard Y., Boileau P. (2011). Neonatal Bacterial Meningitis: 444 Cases in 7 Years. *Pediatr Infect Dis J* **30**, 212-217.

Gaynes R., Edwards J. R., National Nosocomial Infections Surveillance System. (2005). Overview of nosocomial infections caused by gram-negative bacilli. *Clin Infect Dis* **41**, 848-854.

Gbarah A., Gahmberg C. G., Ofek I., Jacobi U., Sharon N. (1991). Identification of the leukocyte adhesion molecules CD11 and CD18 as receptors for type 1-fimbriated (mannose-specific) *Escherichia coli*. *Infect Immun* **59**, 4524-4530.

Ghigo J. M. (2001). Natural conjugative plasmids induce bacterial biofilm development. *Nature* **412**, 442-445.

Girón J. A., Ho A. S., Schoolnik G. K. (1991). An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* **254**, 710-713.

Girón J. A., Torres A. G., Freer E., Kaper J. B. (2002). The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. *Mol Microbiol* **44**, 361-379.

Glode M. P., Sutton A., Moxon E. R., Robbins J. B. (1977). Pathogenesis of neonatal *Escherichia coli* meningitis: induction of bacteremia and meningitis in infant rats fed *E. coli* K1. *Infect Immun* **16**, 75-80.

Goluszko P., Popov V., Selvarangan R., Nowicki S., Pham T., Nowicki B. J. (1997). Dr fimbriae operon of uropathogenic *Escherichia coli* mediate microtubule-dependent invasion to the HeLa epithelial cell line. *J Infect Dis* **176**, 158-167.

Gomez-Duarte O. G., Chattopadhyay S., Weissman S. J., Girón J. A., Kaper J. B., Sokurenko E. V. (2007). Genetic diversity of the gene cluster encoding longus, a type IV pilus of enterotoxigenic *Escherichia coli*. *J Bacteriol* **189**, 9145-9149.

Gophna U., Oelschlaeger T. A., Hacker J., Ron E. Z. (2002). Role of fibronectin in curli-mediated internalization. *FEMS Microbiol Lett* **212**, 55-58.

Gophna U., Barlev M., Seijffers R., Oelschlager T. A., Hacker J., Ron E. Z. (2001). Curli fibers mediate internalization of *Escherichia coli* by eukaryotic cells. *Infect Immun* **69**, 2659-2665.

Göransson M. & Uhlin B. E. (1984). Environmental temperature regulates transcription of a virulence pili operon in *E. coli*. *EMBO J* **3**, 2885-2888.

Göransson M., Sondén B., Nilsson P., Dagberg B., Forsman K., Emanuelsson K., Uhlin B. E. (1990). Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature* **344**, 682-685.

Gordon D. M. & Cowling A. (2003). The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology* **149**, 3575-3586.

Gordon D. M. & FitzGibbon F. (1999). The distribution of enteric bacteria from Australian mammals: host and geographical effects. *Microbiology* **145 (Pt 10)**, 2663-2671.

Gosset G., Zhang Z., Nayyar S., Cuevas W. A., Saier M. H.,Jr. (2004). Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *J Bacteriol* **186**, 3516-3524.

Gottesman S., Trisler P., Torres-Cabassa A. (1985). Regulation of capsular polysaccharide synthesis in *Escherichia coli* K-12: characterization of three regulatory genes. *J Bacteriol* **162**, 1111-1119.

Grainger D. C., Hurd D., Harrison M., Holdstock J., Busby S. J. (2005). Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci U S A* **102**, 17693-17698.

Grozdanov L., Raasch C., Schulze J., Sonnenborn U., Gottschalk G., Hacker J., Dobrindt U. (2004). Analysis of the genome structure of the nonpathogenic probiotic *Escherichia coli* strain Nissle 1917. *J Bacteriol* **186**, 5432-5441.

Guignot J., Peiffer I., Bernet-Camard M. F., Lublin D. M., Carnoy C., Moseley S. L., Servin A. L. (2000). Recruitment of CD55 and CD66e brush border-associated glycosylphosphatidylinositol-anchored proteins by members of the Afa/Dr diffusely adhering family of *Escherichia coli* that infect the human polarized intestinal Caco-2/TC7 cells. *Infect Immun* **68**, 3554-3563.

Guiral E., Bosch J., Vila J., Soto S. M. (2011). Prevalence of *Escherichia coli* among samples collected from the genital tract in pregnant and nonpregnant women: relationship with virulence. *FEMS Microbiol Lett* **314**, 170-173.

Gupta D. S., Jann B., Jann K. (1984). *Escherichia coli* O18ac antigen: structure of the O-specific polysaccharide moiety. *Infect Immun* **45**, 203-209.

Gupta K., Scholes D., Stamm W. E. (1999). Increasing prevalence of antimicrobial resistance among uropathogens causing acute uncomplicated cystitis in women. *JAMA* **281**, 736-738.

Gupta K., Stapleton A. E., Hooton T. M., Roberts P. L., Fennell C. L., Stamm W. E. (1998). Inverse association of H_2O_2-producing lactobacilli and vaginal *Escherichia coli* colonization in women with recurrent urinary tract infections. *J Infect Dis* **178**, 446-450.

Gupte G., Woodward C., Stout V. (1997). Isolation and characterization of *rcsB* mutations that affect colanic acid capsule synthesis in *Escherichia coli* K-12. *J Bacteriol* **179**, 4328-4335.

Gusarov I. & Nudler E. (1999). The mechanism of intrinsic transcription termination. *Mol Cell* **3**, 495-504.

Guttman D. S. & Dykhuizen D. E. (1994). Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* **266**, 1380-1383.

Hagiwara D., Sugiura M., Oshima T., Mori H., Aiba H., Yamashino T., Mizuno T. (2003). Genome-wide analyses revealing a signaling network of the RcsC-YojN-RcsB phosphorelay system in *Escherichia coli*. *J Bacteriol* **185**, 5735-5746.

Hahn E., Wild P., Hermanns U., Sebbel P., Glockshuber R., Häner M., Taschner N., Burkhard P., Aebi U., Müller S. A. (2002). Exploring the 3D molecular architecture of *Escherichia coli* type 1 pili. *J Mol Biol* **323**, 845-857.

Hall-Stoodley L., Costerton J. W., Stoodley P. (2004). Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol* **2**, 95-108.

Hammar M., Arnqvist A., Bian Z., Olsén A., Normark S. (1995). Expression of two *csg* operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. *Mol Microbiol* **18**, 661-670.

Hancock V. & Klemm P. (2007). Global gene expression profiling of asymptomatic bacteriuria *Escherichia coli* during biofilm growth in human urine. *Infect Immun* **75**, 966-976.

Hancock V., Witsø I. L., Klemm P. (2011). Biofilm formation as a function of adhesin, growth medium, substratum and strain type. *Int J Med Microbiol* **301**, 570-576.

Hayashi F., Smith K. D., Ozinsky A., Hawn T. R., Yi E. C., Goodlett D. R., Eng J. K., Akira S., Underhill D. M., Aderem A. (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**, 1099-1103.

Hayashi H., Takahashi R., Nishi T., Sakamoto M., Benno Y. (2005). Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism. *J Med Microbiol* **54**, 1093-1101.

Hendrickx A. P., Budzik J. M., Oh S. Y., Schneewind O. (2011). Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. *Nat Rev Microbiol* **9**, 166-176.

Hernandes R. T., Velsko I., Sampaio S. C., Elias W. P., Robins-Browne R. M., Gomes T. A., Girón J. A. (2011). Fimbrial adhesins produced by atypical enteropathogenic *Escherichia coli* strains. *Appl Environ Microbiol* .

Herzer P. J., Inouye S., Inouye M., Whittam T. S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J Bacteriol* **172**, 6175-6181.

Hinde P., Deighan P., Dorman C. J. (2005). Characterization of the detachable Rho-dependent transcription terminator of the *fimE* gene in *Escherichia coli* K-12. *J Bacteriol* **187**, 8256-8266.

Hoffman S., Sorkin B. C., White P. C., Brackenbury R., Mailhammer R., Rutishauser U., Cunningham B. A., Edelman G. M. (1982). Chemical characterization of a neural cell adhesion molecule purified from embryonic brain membranes. *J Biol Chem* **257**, 7720-7729.

Holden N. J. & Gally D. L. (2004). Switches, cross-talk and memory in *Escherichia coli* adherence. *J Med Microbiol* **53**, 585-593.

Holden N. J., Uhlin B. E., Gally D. L. (2001). PapB paralogues and their effect on the phase variation of type 1 fimbriae in *Escherichia coli*. *Mol Microbiol* **42**, 319-330.

Holden N. J., Totsika M., Mahler E., Roe A. J., Catherwood K., Lindner K., Dobrindt U., Gally D. L. (2006). Demonstration of regulatory cross-talk between P fimbriae and type 1 fimbriae in uropathogenic *Escherichia coli*. *Microbiology* **152**, 1143-1153.

Holmqvist E., Reimegård J., Sterk M., Grantcharova N., Römling U., Wagner E. G. (2010). Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J* **29**, 1840-1850.

Hommais F., Krin E., Laurent-Winter C., Soutourina O., Malpertuy A., Le Caer J. P., Danchin A., Bertin P. (2001). Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* **40**, 20-36.

Hoogkamp-Korstanje J. A., Lindner J. G., Marcelis J. H., den Daas-Slagt H., de Vos N. M. (1979). Composition and ecology of the human intestinal flora. *Antonie Van Leeuwenhoek* **45**, 35-40.

Hooper L. V. & Macpherson A. J. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol* **10**, 159-169.

House B., Kus J. V., Prayitno N., Mair R., Que L., Chingcuanco F., Gannon V., Cvitkovitch D. G., Barnett Foster D. (2009). Acid-stress-induced changes in enterohaemorrhagic *Escherichia coli* O157:H7 virulence. *Microbiology* **155**, 2907-2918.

Houwink A. L. & van Iterson W. (1950). Electron microscopical observations on bacterial cytology; a study on flagellation. *Biochim Biophys Acta* **5**, 10-44.

Huang S. H., Stins M. F., Kim K. S. (2000). Bacterial penetration across the blood-brain barrier during the development of neonatal meningitis. *Microbes Infect* **2**, 1237-1244.

Huang S. H., Wass C., Fu Q., Prasadarao N. V., Stins M., Kim K. S. (1995). *Escherichia coli* invasion of brain microvascular endothelial cells in vitro and in vivo: molecular cloning and characterization of invasion gene *ibe10*. *Infect Immun* **63**, 4470-4475.

Huang Y. H., Ferrières L., Clarke D. J. (2006). The role of the Rcs phosphorelay in *Enterobacteriaceae* . *Res Microbiol* **157**, 206-212.

Hudault S., Guignot J., Servin A. L. (2001). *Escherichia coli* strains colonising the gastrointestinal tract protect germfree mice against *Salmonella typhimurium* infection. *Gut* **49**, 47-55.

Huisman T. T. & de Graaf F. K. (1995). Negative control of *fae* (K88) expression by the 'global' regulator Lrp is modulated by the 'local' regulator FaeA and affected by DNA methylation. *Mol Microbiol* **16**, 943-953.

Hull R. A., Gill R. E., Hsu P., Minshew B. H., Falkow S. (1981). Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. *Infect Immun* **33**, 933-938.

Hung C. S., Bouckaert J., Hung D., Pinkner J., Widberg C., DeFusco A., Auguste C. G., Strouse R., Langermann S., Waksman G., Hultgren S. J. (2002). Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol Microbiol* **44**, 903-915.

Huys G., Cnockaert M., Janda J. M., Swings J. (2003). *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. *Int J Syst Evol Microbiol* **53**, 807-810.

Ingle D. J., Clermont O., Skurnik D., Denamur E., Walk S. T., Gordon D. M. (2011). Biofilm formation by and thermal niche and virulence characteristics of *Escherichia* spp. *Appl Environ Microbiol* **77**, 2695-2700.

Isaacson R. E., Fusco P. C., Brinton C. C., Moon H. W. (1978). In vitro adhesion of *Escherichia coli* to porcine small intestinal epithelial cells: pili as adhesive factors. *Infect Immun* **21**, 392-397.

Ishihama A. (2010). Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiol Rev* **34**, 628-645.

Ishii S., Ksoll W. B., Hicks R. E., Sadowsky M. J. (2006). Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. *Appl Environ Microbiol* **72**, 612-621.

Ishizuka H., Hanamura A., Kunimura T., Aiba H. (1993). A lowered concentration of cAMP receptor protein caused by glucose is an important determinant for catabolite repression in *Escherichia coli*. *Mol Microbiol* **10**, 341-350.

Jacobs A. A., Simons B. H., de Graaf F. K. (1987). The role of lysine-132 and arginine-136 in the receptor-binding domain of the K99 fibrillar subunit. *EMBO J* **6**, 1805-1808.

Jansson L., Tobias J., Lebens M., Svennerholm A. M., Teneberg S. (2006). The major subunit, CfaB, of colonization factor antigen i from enterotoxigenic *Escherichia coli* is a glycosphingolipid binding protein. *Infect Immun* **74**, 3488-3497.

Johnson J. R., Oswald E., O'Bryan T. T., Kuskowski M. A., Spanjaard L. (2002). Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands. *J Infect Dis* **185**, 774-784.

Jones C. H., Pinkner J. S., Roth R., Heuser J., Nicholes A. V., Abraham S. N., Hultgren S. J. (1995). FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the *Enterobacteriaceae*. *Proc Natl Acad Sci U S A* **92**, 2081-2085.

Jordi B. J., op den Camp I. E., de Haan L. A., van der Zeijst B. A., Gaastra W. (1993). Differential decay of RNA of the CFA/I fimbrial operon and control of relative gene expression. *J Bacteriol* **175**, 7976-7981.

Joyce S. A. & Dorman C. J. (2002). A Rho-dependent phase-variable transcription terminator controls expression of the FimE recombinase in *Escherichia coli*. *Mol Microbiol* **45**, 1107-1117.

Kai-Larsen Y., Lüthje P., Chromek M., Peters V., Wang X., Holm Å., Kádas L., Hedlund K. O., Johansson J.& other authors. (2010). Uropathogenic *Escherichia coli* modulates immune responses and its curli fimbriae interact with the antimicrobial peptide LL-37. *PLoS Pathog* **6**, e1001010.

Kailasan Vanaja S., Bergholz T. M., Whittam T. S. (2009). Characterization of the *Escherichia coli* O157:H7 Sakai GadE regulon. *J Bacteriol* **191**, 1868-1877.

Kalir S., McClure J., Pabbaraju K., Southward C., Ronen M., Leibler S., Surette M. G., Alon U. (2001). Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science* **292**, 2080-2083.

Kaper J. B., Nataro J. P., Mobley H. L. (2004). Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**, 123-140.

Kapust R. B. & Waugh D. S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668-1674.

Karatan E. & Watnick P. (2009). Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiol Mol Biol Rev* **73**, 310-347.

Kauffmann F. (1947). The serology of the coli group. *J Immunol* **57**, 71-100.

Kawula T. H. & Orndorff P. E. (1991). Rapid site-specific DNA inversion in *Escherichia coli* mutants lacking the histonelike protein H-NS. *J Bacteriol* **173**, 4116-4123.

Khan M. A. & Isaacson R. E. (2002). Identification of *Escherichia coli* genes that are specifically expressed in a murine model of septicemic infection. *Infect Immun* **70**, 3404-3412.

Kido M., Yamanaka K., Mitani T., Niki H., Ogura T., Hiraga S. (1996). RNase E polypeptides lacking a carboxyl-terminal half suppress a *mukB* mutation in *Escherichia coli*. *J Bacteriol* **178**, 3917-3925.

Kikuchi T., Mizunoe Y., Takade A., Naito S., Yoshida S. (2005). Curli fibers are required for development of biofilm architecture in *Escherichia coli* K-12 and enhance bacterial adherence to human uroepithelial cells. *Microbiol Immunol* **49**, 875-884.

Kim J., Adhya S., Garges S. (1992). Allosteric changes in the cAMP receptor protein of *Escherichia coli*: hinge reorientation. *Proc Natl Acad Sci U S A* **89**, 9700-9704.

Kim K. J., Elliott S. J., Di Cello F., Stins M. F., Kim K. S. (2003). The K1 capsule modulates trafficking of *E. coli*-containing vacuoles and enhances intracellular bacterial survival in human brain microvascular endothelial cells. *Cell Microbiol* **5**, 245-252.

Kim K. S. (2008). Mechanisms of microbial traversal of the blood-brain barrier. *Nat Rev Microbiol* **6**, 625-634.

Kim K. S. (2003). Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury. *Nat Rev Neurosci* **4**, 376-385.

Kim K. S., Itabashi H., Gemski P., Sadoff J., Warren R. L., Cross A. S. (1992). The K1 capsule is the critical determinant in the development of *Escherichia coli* meningitis in the rat. *J Clin Invest* **90**, 897-905.

Klemm P., Hancock V., and Schembri M. A. (2010). Fimbrial adhesins from extraintestinal *Escherichia coli*. *Environ Microbiol Rep* **2**, 628-640.

Klemm P. (1986). Two regulatory *fim* genes, *fimB* and *fimE*, control the phase variation of type 1 fimbriae in *Escherichia coli*. *EMBO J* **5**, 1389-1393.

Klemm P. & Christiansen G. (1987). Three fim genes required for the regulation of length and mediation of adhesion of *Escherichia coli* type 1 fimbriae. *Mol Gen Genet* **208**, 439-445.

Kline K. A., Dodson K. W., Caparon M. G., Hultgren S. J. (2010). A tale of two pili: assembly and function of pili in bacteria. *Trends Microbiol* **18**, 224-232.

Ko M. & Park C. (2000). Two novel flagellar components and H-NS are involved in the motor function of *Escherichia coli* . *J Mol Biol* **303**, 371-382.

Komeda Y. (1986). Transcriptional control of flagellar genes in *Escherichia coli* K-12. *J Bacteriol* **168**, 1315-1318.

Koo J. T., Choe J., Moseley S. L. (2004). HrpA, a DEAH-box RNA helicase, is involved in mRNA processing of a fimbrial operon in *Escherichia coli*. *Mol Microbiol* **52**, 1813-1826.

Korea C. G., Ghigo J. M., Beloin C. (2011). The sweet connection: Solving the riddle of multiple sugar-binding fimbrial adhesins in *Escherichia coli*: Multiple *E. coli* fimbriae form a versatile arsenal of sugar-binding lectins potentially involved in surface-colonisation and tissue tropism. *Bioessays* **33**, 300-311.

Korea C. G., Badouraly R., Prevost M. C., Ghigo J. M., Beloin C. (2010). *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. *Environ Microbiol* **12**, 1957-1977.

Korhonen T. K. & Finne J. (1985). Agglutination assays for detecting bacterial binding specificities. In *Enterobacterial Surface Antigens: Methods for Molecular Characterization.* pp. 301-303. Edited by T. K. Korhonen, E. A. Dawes & P. H. Mäkelä. Amsterdam, The Netherlands.: Elsevier Science Publishers.

Korhonen T. K., Virkola R., Holthöfer H. (1986). Localization of binding sites for purified *Escherichia coli* P fimbriae in the human kidney. *Infect Immun* **54**, 328-332.

Korhonen T. K., Väisänen V., Saxén H., Hultberg H., Svenson S. B. (1982). P-antigen-recognizing fimbriae from human uropathogenic *Escherichia coli* strains. *Infect Immun* **37**, 286-291.

Korhonen T. K., Väisänen-Rhen V., Rhen M., Pere A., Parkkinen J., Finne J. (1984). *Escherichia coli* fimbriae recognizing sialyl galactosides. *J Bacteriol* **159**, 762-766.

Korhonen T. K., Parkkinen J., Hacker J., Finne J., Pere A., Rhen M., Holthöfer H. (1986). Binding of *Escherichia coli* S fimbriae to human kidney epithelium. *Infect Immun* **54**, 322-327.

Korhonen T. K., Valtonen M. V., Parkkinen J., Väisänen-Rhen V., Finne J., Ørskov F., Ørskov I., Svenson S. B., Mäkelä P. H. (1985). Serotypes, hemolysin production, and receptor recognition of *Escherichia coli* strains associated with neonatal sepsis and meningitis. *Infect Immun* **48**, 486-491.

Korotkova N., Chattopadhyay S., Tabata T. A., Beskhlebnaya V., Vigdorovich V., Kaiser B. K., Strong R. K., Dykhuizen D. E., Sokurenko E. V., Moseley S. L. (2007). Selection for functional diversity drives accumulation of point mutations in Dr adhesins of *Escherichia coli*. *Mol Microbiol* **64**, 180-194.

Krogfelt K. A., Bergmans H., Klemm P. (1990). Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* type 1 fimbriae. *Infect Immun* **58**, 1995-1998.

Krohn M. A., Thwin S. S., Rabe L. K., Brown Z., Hillier S. L. (1997). Vaginal colonization by Escherichia coli as a risk factor for very low birth weight delivery and other perinatal complications. *J Infect Dis* **175**, 606-610.

Kuehn M. J., Heuser J., Normark S., Hultgren S. J. (1992). P pili in uropathogenic *E. coli* are composite fibres with distinct fibrillar adhesive tips. *Nature* **356**, 252-255.

Kukkonen M., Saarela S., Lähteenmäki K., Hynönen U., Westerlund-Wikström B., Rhen M., Korhonen T. K. (1998). Identification of two laminin-binding fimbriae, the type 1 fimbria of *Salmonella enterica* serovar typhimurium and the G fimbria of *Escherichia coli*, as plasminogen receptors. *Infect Immun* **66**, 4965-4970.

Kukkonen M., Raunio T., Virkola R., Lähteenmäki K., Mäkelä P. H., Klemm P., Clegg S., Korhonen T. K. (1993). Basement membrane carbohydrate as a target for bacterial adhesion: binding of type I fimbriae of *Salmonella enterica* and *Escherichia coli* to laminin. *Mol Microbiol* **7**, 229-237.

Kutsukake K., Ohya Y., Iino T. (1990). Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. *J Bacteriol* **172**, 741-747.

Kvist M., Hancock V., Klemm P. (2008). Inactivation of efflux pumps abolishes bacterial biofilm formation. *Appl Environ Microbiol* **74**, 7376-7382.

Lacher D. W., Steinsland H., Blank T. E., Donnenberg M. S., Whittam T. S. (2007). Molecular evolution of typical enteropathogenic *Escherichia coli*: clonal analysis by multilocus sequence typing and virulence gene allelic profiling. *J Bacteriol* **189**, 342-350.

Lähteenmäki K., Kuusela P., Korhonen T. K. (2001). Bacterial plasminogen activators and receptors. *FEMS Microbiol Rev* **25**, 531-552.

Lan R., Alles M. C., Donohoe K., Martinez M. B., Reeves P. R. (2004). Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect Immun* **72**, 5080-5088.

Lane M. C., Simms A. N., Mobley H. L. (2007a). Complex interplay between type 1 fimbrial expression and flagellum-mediated motility of uropathogenic *Escherichia coli*. *J Bacteriol* **189**, 5523-5533.

Lane M. C., Alteri C. J., Smith S. N., Mobley H. L. (2007b). Expression of flagella is coincident with uropathogenic *Escherichia coli* ascension to the upper urinary tract. *Proc Natl Acad Sci U S A* **104**, 16669-16674.

Lane M. C., Lockatell V., Monterosso G., Lamphier D., Weinert J., Hebel J. R., Johnson D. E., Mobley H. L. (2005). Role of motility in the colonization of uropathogenic *Escherichia coli* in the urinary tract. *Infect Immun* **73**, 7644-7656.

Lasaro M. A., Salinger N., Zhang J., Wang Y., Zhong Z., Goulian M., Zhu J. (2009). F1C fimbriae play an important role in biofilm formation and intestinal colonization by the *Escherichia coli* commensal strain Nissle 1917. *Appl Environ Microbiol* **75**, 246-251.

Lawrence J. G. & Ochman H. (1998). Molecular archaeology of the *Escherichia coli* genome. *Proc Natl Acad Sci U S A* **95**, 9413-9417.

Le Trong I., Aprikian P., Kidd B. A., Forero-Shelton M., Tchesnokova V., Rajagopal P., Rodriguez V., Interlandi G., Klevit R.& other authors. (2010). Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like β sheet twisting. *Cell* **141**, 645-655.

Leatham M. P., Banerjee S., Autieri S. M., Mercado-Lubo R., Conway T., Cohen P. S. (2009). Precolonized human commensal *Escherichia coli* strains serve as a barrier to *E. coli* O157:H7 growth in the streptomycin-treated mouse intestine. *Infect Immun* **77**, 2876-2886.

Leathart J. B. & Gally D. L. (1998). Regulation of type 1 fimbrial expression in uropathogenic *Escherichia coli*: heterogeneity of expression through sequence changes in the *fim* switch region. *Mol Microbiol* **28**, 371-381.

Lecointre G., Rachdi L., Darlu P., Denamur E. (1998). *Escherichia coli* molecular phylogeny using the incongruence length difference test. *Mol Biol Evol* **15**, 1685-1695.

Leffler H. & Svanborg-Edén C. (1981). Glycolipid receptors for uropathogenic *Escherichia coli* on human erythrocytes and uroepithelial cells. *Infect Immun* **34**, 920-929.

Lehti T. A., Bauchart P., Heikkinen J., Hacker J., Korhonen T. K., Dobrindt U., Westerlund-Wikström B. (2010). Mat fimbriae promote biofilm formation by meningitis-associated *Escherichia coli*. *Microbiology* **156**, 2408-2417.

Levine M. M., Bergquist E. J., Nalin D. R., Waterman D. H., Hornick R. B., Young C. R., Sotman S. (1978). *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet* **1**, 1119-1122.

Ley R. E., Hamady M., Lozupone C., Turnbaugh P. J., Ramey R. R., Bircher J. S., Schlegel M. L., Tucker T. A., Schrenzel M. D., Knight R., Gordon J. I. (2008). Evolution of mammals and their gut microbes. *Science* **320**, 1647-1651.

Li K., Zhou W., Hong Y., Sacks S. H., Sheerin N. S. (2009). Synergy between type 1 fimbriae expression and C3 opsonisation increases internalisation of *E. coli* by human tubular epithelial cells. *BMC Microbiol* **9**, 64.

Li X., Rasko D. A., Lockatell C. V., Johnson D. E., Mobley H. L. (2001). Repression of bacterial motility by a novel fimbrial gene product. *EMBO J* **20**, 4854-4862.

Liang S. T., Dennis P. P., Bremer H. (1998). Expression of *lacZ* from the promoter of the *Escherichia coli spc* operon cloned into vectors carrying the W205 *trp-lac* fusion. *J Bacteriol* **180**, 6090-6100.

Lindberg F., Lund B., Johansson L., Normark S. (1987). Localization of the receptor-binding protein adhesin at the tip of the bacterial pilus. *Nature* **328**, 84-87.

Lindberg S., Xia Y., Sondén B., Göransson M., Hacker J., Uhlin B. E. (2008). Regulatory Interactions among adhesin gene systems of uropathogenic *Escherichia coli*. *Infect Immun* **76**, 771-780.

Lindberg U., Hanson L. A., Jodal U., Lidin-Janson G., Lincoln K., Olling S. (1975). Asymptomatic bacteriuria in schoolgirls. II. Differences in *Escherichia coli* causing asymptomatic bacteriuria. *Acta Paediatr Scand* **64**, 432-436.

Link A. J., Phillips D., Church G. M. (1997). Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* **179**, 6228-6237.

Liu X. & Matsumura P. (1994). The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. *J Bacteriol* **176**, 7345-7351.

Lomovskaya O., Lewis K., Matin A. (1995). EmrR is a negative regulator of the *Escherichia coli* multidrug resistance pump EmrAB. *J Bacteriol* **177**, 2328-2334.

Lopez P. J., Marchand I., Joyce S. A., Dreyfus M. (1999). The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing *in vivo*. *Mol Microbiol* **33**, 188-199.

Low A. S., Holden N., Rosser T., Roe A. J., Constantinidou C., Hobman J. L., Smith D. G., Low J. C., Gally D. L. (2006). Analysis of fimbrial gene clusters and their expression in enterohaemorrhagic *Escherichia coli* O157:H7. *Environ Microbiol* **8**, 1033-1047.

Luck S. N., Badea L., Bennett-Wood V., Robins-Browne R., Hartland E. L. (2006). Contribution of FliC to epithelial cell invasion by enterohemorrhagic *Escherichia coli* O113:H21. *Infect Immun* **74**, 6999-7004.

Lukjancenko O., Wassenaar T. M., Ussery D. W. (2010). Comparison of 61 sequenced *Escherichia coli* genomes. *Microb Ecol* **60**, 708-720.

Lund B., Lindberg F., Marklund B. I., Normark S. (1987). The PapG protein is the α-D-galactopyranosyl-(1→4)-β-D-galactopyranose-binding adhesin of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A* **84**, 5898-5902.

Luo C., Walk S. T., Gordon D. M., Feldgarden M., Tiedje J. M., Konstantinidis K. T. (2011). Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Proc Natl Acad Sci U S A* **108**, 7200-7205.

Macfarlane S. & Macfarlane G. T. (2006). Composition and metabolic activities of bacterial biofilms colonizing food residues in the human gut. *Appl Environ Microbiol* **72**, 6204-6211.

Macfarlane S., Bahrami B., Macfarlane G. T. (2011). Mucosal biofilm communities in the human intestinal tract. *Adv Appl Microbiol* **75**, 111-143.

Mahajan A., Currie C. G., Mackie S., Tree J., McAteer S., McKendrick I., McNeilly T. N., Roe A., La Ragione R. M.& other authors. (2009). An investigation of the expression and adhesin function of H7 flagella in the interaction of *Escherichia coli* O157 : H7 with bovine intestinal epithelium. *Cell Microbiol* **11**, 121-137.

Majdalani N. & Gottesman S. (2005). The Rcs phosphorelay: a complex signal transduction system. *Annu Rev Microbiol* **59**, 379-405.

Majdalani N., Hernandez D., Gottesman S. (2002). Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol* **46**, 813-826.

Majdalani N., Heck M., Stout V., Gottesman S. (2005). Role of RcsF in signaling to the Rcs phosphorelay pathway in *Escherichia coli*. *J Bacteriol* **187**, 6770-6778.

Majdalani N., Cunning C., Sledjeski D., Elliott T., Gottesman S. (1998). DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95**, 12462-12467.

Mangan S. & Alon U. (2003). Structure and function of the feed-forward loop network motif. *Proc Natl Acad Sci U S A* **100**, 11980-11985.

Martin C. (1996). The *clp* (CS31A) operon is negatively controlled by Lrp, ClpB, and L-alanine at the transcriptional level. *Mol Microbiol* **21**, 281-292.

Martin F. H. & Tinoco I.,Jr. (1980). DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Res* **8**, 2295-2299.

Martinez J. J., Mulvey M. A., Schilling J. D., Pinkner J. S., Hultgren S. J. (2000). Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J* **19**, 2803-2812.

Matsumura K., Furukawa S., Ogihara H., Morinaga Y. (2011). Roles of multidrug efflux pumps on the biofilm formation of *Escherichia coli* K-12. *Biocontrol Sci* **16**, 69-72.

Maurelli A. T., Fernández R. E., Bloch C. A., Rode C. K., Fasano A. (1998). "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc Natl Acad Sci U S A* **95**, 3943-3948.

Mazariego-Espinosa K., Cruz A., Ledesma M. A., Ochoa S. A., Xicohtencatl-Cortes J. (2010). Longus, a type IV pilus of enterotoxigenic *Escherichia coli*, is involved in adherence to intestinal epithelial cells. *J Bacteriol* **192**, 2791-2800.

McBroom A. J. & Kuehn M. J. (2007). Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. *Mol Microbiol* **63**, 545-558.

McClain M. S., Blomfield I. C., Eisenstein B. I. (1991). Roles of *fimB* and *fimE* in site-specific DNA inversion associated with phase variation of type 1 fimbriae in *Escherichia coli*. *J Bacteriol* **173**, 5308-5314.

McDougald D., Rice S. A., Barraud N., Steinberg P. D., Kjelleberg S. (2012). Should we stay or should we go: mechanisms and ecological consequences for biofilm dispersal. *Nat Rev Microbiol* **10**, 39-50.

McDowall K. J., Lin-Chao S., Cohen S. N. (1994). A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J Biol Chem* **269**, 10790-10796.

McDowall K. J., Hernandez R. G., Lin-Chao S., Cohen S. N. (1993). The *ams-1* and *rne-3071* temperature-sensitive mutations in the *ams* gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the *Escherichia coli mre* locus. *J Bacteriol* **175**, 4245-4249.

Melican K., Sandoval R. M., Kader A., Josefsson L., Tanner G. A., Molitoris B. A., Richter-Dahlfors A. (2011). Uropathogenic *Escherichia coli* P and Type 1 fimbriae act in synergy in a living host to facilitate renal colonization leading to nephron obstruction. *PLoS Pathog* **7**, e1001298.

Mellmann A., Harmsen D., Cummings C. A., Zentz E. B., Leopold S. R., Rico A., Prior K., Szczepanowski R., Ji Y.& other authors. (2011). Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One* **6**, e22751.

Mendoza-Vargas A., Olvera L., Olvera M., Grande R., Vega-Alvarado L., Taboada B., Jimenez-Jacinto V., Salgado H., Juárez K.& other authors. (2009). Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One* **4**, e7526.

Miller V. L. & Mekalanos J. J. (1988). A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR* . *J Bacteriol* **170**, 2575-2583.

Mittal R. & Prasadarao N. V. (2008). Outer membrane protein A expression in *Escherichia coli* K1 is required to prevent the maturation of myeloid dendritic cells and the induction of IL-10 and TGF-β. *J Immunol* **181**, 2672-2682.

Mobley H. L., Green D. M., Trifillis A. L., Johnson D. E., Chippendale G. R., Lockatell C. V., Jones B. D., Warren J. W. (1990). Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun* **58**, 1281-1289.

Mobley H. L., Jarvis K. G., Elwood J. P., Whittle D. I., Lockatell C. V., Russell R. G., Johnson D. E., Donnenberg M. S., Warren J. W. (1993). Isogenic P-fimbrial deletion mutants of pyelonephritogenic *Escherichia coli*: the role of αGal(1-4)βGal binding in virulence of a wild-type strain. *Mol Microbiol* **10**, 143-155.

Moch T., Hoschützky H., Hacker J., Kröncke K. D., Jann K. (1987). Isolation and characterization of the α-sialyl- β-2,3-galactosyl-specific adhesin from fimbriated *Escherichia coli*. *Proc Natl Acad Sci U S A* **84**, 3462-3466.

Monds R. D. & O'Toole G. A. (2009). The developmental model of microbial biofilms: ten years of a paradigm up for review. *Trends Microbiol* **17**, 73-87.

Moreno E., Johnson J. R., Pérez T., Prats G., Kuskowski M. A., Andreu A. (2009). Structure and urovirulence characteristics of the fecal *Escherichia coli* population among healthy women. *Microbes Infect* **11**, 274-280.

Moriel D. G., Bertoldi I., Spagnuolo A., Marchi S., Rosini R., Nesta B., Pastorello I., Corea V. A., Torricelli G.& other authors. (2010). Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A* **107**, 9072-9077.

Morschhäuser J., Uhlin B. E., Hacker J. (1993). Transcriptional analysis and regulation of the *sfa* determinant coding for S fimbriae of pathogenic *Escherichia coli* strains. *Mol Gen Genet* **238**, 97-105.

Moulin-Schouleur M., Répérant M., Laurent S., Brée A., Mignon-Grasteau S., Germon P., Rasschaert D., Schouler C. (2007). Extraintestinal pathogenic *Escherichia coli* strains of avian and human origin: link between phylogenetic relationships and common virulence patterns. *J Clin Microbiol* **45**, 3366-3376.

Moulin-Schouleur M., Schouler C., Tailliez P., Kao M. R., Brée A., Germon P., Oswald E., Mainil J., Blanco M., Blanco J. (2006). Common virulence factors and genetic relationships between O18:K1:H7 *Escherichia coli* isolates of human and avian origin. *J Clin Microbiol* **44**, 3484-3492.

Moxon E. R. & Ostrow P. T. (1977). *Haemophilus influenzae* meningitis in infant rats: role of bacteremia in pathogenesis of age-dependent inflammatory responses in cerebrospinal fluid. *J Infect Dis* **135**, 303-307.

Müller C. M., Åberg A., Straseviçiene J., Emödy L., Uhlin B. E., Balsalobre C. (2009). Type 1 fimbriae, a colonization factor of uropathogenic *Escherichia coli*, are controlled by the metabolic sensor CRP-cAMP. *PLoS Pathog* **5**, e1000303.

Müller C. M., Dobrindt U., Nagy G., Emödy L., Uhlin B. E., Hacker J. (2006). Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic *Escherichia coli*. *J Bacteriol* **188**, 5428-5438.

Müsken M., Di Fiore S., Dötsch A., Fischer R., Häussler S. (2010). Genetic determinants of *Pseudomonas aeruginosa* biofilm establishment. *Microbiology* **156**, 431-441.

Nagy G., Dobrindt U., Grozdanov L., Hacker J., Emõdy L. (2005). Transcriptional regulation through RfaH contributes to intestinal colonization by *Escherichia coli*. *FEMS Microbiol Lett* **244**, 173-180.

Narciso A., Nunes F., Amores T., Lito L., Melo-Cristino J., Duarte A. (2011). Persistence of uropathogenic *Escherichia coli* strains in the host for long periods of time: relationship between phylogenetic groups and virulence factors. *Eur J Clin Microbiol Infect Dis* .

Natale P., Brüser T., Driessen A. J. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—Distinct translocases and mechanisms. *Biochim Biophys Acta* **1778**, 1735-1756.

Nègre V. L., Bonacorsi S., Schubert S., Bidet P., Nassif X., Bingen E. (2004). The siderophore receptor IroN, but not the high-pathogenicity island or the hemin receptor ChuA, contributes to the bacteremic step of *Escherichia coli* neonatal meningitis. *Infect Immun* **72**, 1216-1220.

Niba E. T., Naka Y., Nagase M., Mori H., Kitakawa M. (2007). A genome-wide approach to identify the genes involved in biofilm formation in *E. coli*. *DNA Res* **14**, 237-246.

Nilsson L. M., Thomas W. E., Trintchina E., Vogel V., Sokurenko E. V. (2006). Catch bond-mediated adhesion without a shear threshold: trimannose versus monomannose interactions with the FimH adhesin of *Escherichia coli*. *J Biol Chem* **281**, 16656-16663.

Nilsson P. & Uhlin B. E. (1991). Differential decay of a polycistronic *Escherichia coli* transcript is initiated by RNaseE-dependent endonucleolytic processing. *Mol Microbiol* **5**, 1791-1799.

Nilsson P., Naureckiene S., Uhlin B. E. (1996). Mutations affecting mRNA processing and fimbrial biogenesis in the *Escherichia coli pap* operon. *J Bacteriol* **178**, 683-690.

Nougayrède J. P., Homburg S., Taieb F., Boury M., Brzuszkiewicz E., Gottschalk G., Buchrieser C., Hacker J., Dobrindt U., Oswald E. (2006). *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **313**, 848-851.

Nowicki B., Selvarangan R., Nowicki S. (2001). Family of *Escherichia coli* Dr adhesins: decay-accelerating factor receptor recognition and invasiveness. *J Infect Dis* **183 Suppl 1**, S24-7.

Nowicki B., Moulds J., Hull R., Hull S. (1988). A hemagglutinin of uropathogenic *Escherichia coli* recognizes the Dr blood group antigen. *Infect Immun* **56**, 1057-1060.

Nowicki B., Hart A., Coyne K. E., Lublin D. M., Nowicki S. (1993). Short consensus repeat-3 domain of recombinant decay-accelerating factor is recognized by *Escherichia coli* recombinant Dr adhesin in a model of a cell-cell interaction. *J Exp Med* **178**, 2115-2121.

Nowicki B., Rhen M., Väisänen-Rhen V., Pere A., Korhonen T. K. (1984). Immunofluorescence study of fimbrial phase variation in *Escherichia coli* KS71. *J Bacteriol* **160**, 691-695.

Nowrouzian F. L., Wold A. E., Adlerberth I. (2005). *Escherichia coli* strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. *J Infect Dis* **191**, 1078-1083.

Nuccio S. P. & Bäumler A. J. (2007). Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol Mol Biol Rev* **71**, 551-575.

Obata-Yasuoka M., Ba-Thein W., Tsukamoto T., Yoshikawa H., Hayashi H. (2002). Vaginal *Escherichia coli* share common virulence factor profiles, serotypes and phylogeny with other extraintestinal *E. coli*. *Microbiology* **148**, 2745-2752.

Ochman H. & Wilson A. C. (1987). Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J Mol Evol* **26**, 74-86.

Ochman H. & Selander R. K. (1984). Evidence for clonal population structure in *Escherichia coli*. *Proc Natl Acad Sci U S A* **81**, 198-201.

Ogasawara H., Yamamoto K., Ishihama A. (2011). Role of the biofilm master regulator CsgD in cross-regulation between biofilm formation and flagellar synthesis. *J Bacteriol* **193**, 2587-2597.

Ogura Y., Ooka T., Iguchi A., Toh H., Asadulghani M., Oshima K., Kodama T., Abe H., Nakayama K.& other authors. (2009). Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc Natl Acad Sci U S A* **106**, 17939-17944.

O'Hara A. M. & Shanahan F. (2006). The gut flora as a forgotten organ. *EMBO Rep* **7**, 688-693.

Olsén A., Jonsson A., Normark S. (1989). Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. *Nature* **338**, 652-655.

Olsén A., Arnqvist A., Hammar M., Sukupolvi S., Normark S. (1993). The RpoS sigma factor relieves H-NS-mediated transcriptional repression of *csgA*, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Mol Microbiol* **7**, 523-536.

Orndorff P. E., Devapali A., Palestrant S., Wyse A., Everett M. L., Bollinger R. R., Parker W. (2004). Immunoglobulin-mediated agglutination of and biofilm formation by *Escherichia coli* K-12 require the type 1 pilus fiber. *Infect Immun* **72**, 1929-1938.

Ørskov F. & Ørskov I. (1992). *Escherichia coli* serotyping and disease in man and animals. *Can J Microbiol* **38**, 699-704.

Ørskov F., Sharma V., Ørskov I. (1984). Influence of growth temperature on the development of *Escherichia coli* polysaccharide K antigens. *J Gen Microbiol* **130**, 2681-2684.

O'Ryan M., Prado V., Pickering L. K. (2005). A millennium update on pediatric diarrheal illness in the developing world. *Semin Pediatr Infect Dis* **16**, 125-136.

Oshima T., Ishikawa S., Kurokawa K., Aiba H., Ogasawara N. (2006). *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res* **13**, 141-153.

Oshima T., Aiba H., Masuda Y., Kanaya S., Sugiura M., Wanner B. L., Mori H., Mizuno T. (2002). Transcriptome analysis of all two-component regulatory system mutants of *Escherichia coli* K-12. *Mol Microbiol* **46**, 281-291.

O'Toole G., Kaplan H. B., Kolter R. (2000). Biofilm formation as microbial development. *Annu Rev Microbiol* **54**, 49-79.

Palestrant D., Holzknecht Z. E., Collins B. H., Parker W., Miller S. E., Bollinger R. R. (2004). Microbial biofilms in the gut: visualization by electron microscopy and by acridine orange staining. *Ultrastruct Pathol* **28**, 23-27.

Palmer C., Bik E. M., DiGiulio D. B., Relman D. A., Brown P. O. (2007). Development of the human infant intestinal microbiota. *PLoS Biol* **5**, e177.

Parkkinen J. & Korhonen T. K. (1989). Binding of plasminogen to *Escherichia coli* adhesion proteins. *FEBS Lett* **250**, 437-440.

Parkkinen J., Hacker J., Korhonen T. K. (1991). Enhancement of tissue plasminogen activator-catalyzed plasminogen activation by *Escherichia coli* S fimbriae associated with neonatal septicaemia and meningitis. *Thromb Haemost* **65**, 483-486.

Parkkinen J., Virkola R., Korhonen T. K. (1988a). Identification of factors in human urine that inhibit the binding of *Escherichia coli* adhesins. *Infect Immun* **56**, 2623-2630.

Parkkinen J., Korhonen T. K., Pere A., Hacker J., Soinila S. (1988b). Binding sites in the rat brain for *Escherichia coli* S fimbriae associated with neonatal meningitis. *J Clin Invest* **81**, 860-865.

Parkkinen J., Rogers G. N., Korhonen T., Dahr W., Finne J. (1986). Identification of the O-linked sialyloligosaccharides of glycophorin A as the erythrocyte receptors for S-fimbriated *Escherichia coli*. *Infect Immun* **54**, 37-42.

Parthasarathy G., Yao Y., Kim K. S. (2007). Flagella promote *Escherichia coli* K1 association with and invasion of human brain microvascular endothelial cells. *Infect Immun* **75**, 2937-2945.

Passner J. M. & Steitz T. A. (1997). The structure of a CAP-DNA complex having two cAMP molecules bound to each monomer. *Proc Natl Acad Sci U S A* **94**, 2843-2847.

Paul K., Carlquist W. C., Blair D. F. (2011). Adjusting the spokes of the flagellar motor with the DNA-binding protein H-NS. *J Bacteriol* **193**, 5914-5922.

Pearson M. M. & Mobley H. L. (2008). Repression of motility during fimbrial expression: identification of 14 *mrpJ* gene paralogues in *Proteus mirabilis*. *Mol Microbiol* **69**, 548-558.

Peek A. S., Souza V., Eguiarte L. E., Gaut B. S. (2001). The interaction of protein structure, selection, and recombination on the evolution of the type-1 fimbrial major subunit (*fimA*) from *Escherichia coli*. *J Mol Evol* **52**, 193-204.

Penders J., Thijs C., Vink C., Stelma F. F., Snijders B., Kummeling I., van den Brandt P. A., Stobberingh E. E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**, 511-521.

Pesavento C., Becker G., Sommerfeldt N., Possling A., Tschowri N., Mehlis A., Hengge R. (2008). Inverse regulatory coordination of motility and curli-mediated adhesion in *Escherichia coli*. *Genes Dev* **22**, 2434-2446.

Pfeiffer V., Papenfort K., Lucchini S., Hinton J. C., Vogel J. (2009). Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* **16**, 840-846.

Picard B., Garcia J. S., Gouriou S., Duriez P., Brahimi N., Bingen E., Elion J., Denamur E. (1999). The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* **67**, 546-553.

Pichon C., Héchard C., du Merle L., Chaudray C., Bonne I., Guadagnini S., Vandewalle A., Le Bouguénec C. (2009). Uropathogenic *Escherichia coli* AL511 requires flagellum to enter renal collecting duct cells. *Cell Microbiol* **11**, 616-628.

Pieroni P., Worobec E. A., Paranchych W., Armstrong G. D. (1988). Identification of a human erythrocyte receptor for colonization factor antigen I pili expressed by H10407 enterotoxigenic *Escherichia coli*. *Infect Immun* **56**, 1334-1340.

Pinkner J. S., Remaut H., Buelens F., Miller E., Åberg V., Pemberton N., Hedenström M., Larsson A., Seed P.& other authors. (2006). Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria. *Proc Natl Acad Sci U S A* **103**, 17897-17902.

Pluschke G., Mercer A., Kusećek B., Pohl A., Achtman M. (1983). Induction of bacteremia in newborn rats by *Escherichia coli* K1 is correlated with only certain O (lipopolysaccharide) antigen types. *Infect Immun* **39**, 599-608.

Pouttu R., Puustinen T., Virkola R., Hacker J., Klemm P., Korhonen T. K. (1999). Amino acid residue Ala-62 in the FimH fimbrial adhesin is critical for the adhesiveness of meningitis-associated *Escherichia coli* to collagens. *Mol Microbiol* **31**, 1747-1757.

Pouttu R., Westerlund-Wikström B., Lång H., Alsti K., Virkola R., Saarela U., Siitonen A., Kalkkinen N., Korhonen T. K. (2001). *matB*, a common fimbrillin gene of *Escherichia coli*, expressed in a genetically conserved, virulent clonal group. *J Bacteriol* **183**, 4727-4736.

Power M. L., Littlefield-Wyer J., Gordon D. M., Veal D. A., Slade M. B. (2005). Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. *Environ Microbiol* **7**, 631-640.

Prasadarao N. V. (2002). Identification of *Escherichia coli* outer membrane protein A receptor on human brain microvascular endothelial cells. *Infect Immun* **70**, 4556-4563.

Prasadarao N. V., Wass C. A., Kim K. S. (1996). Endothelial cell GlcNAcβ1-4GlcNAc epitopes for outer membrane protein A enhance traversal of *Escherichia coli* across the blood-brain barrier. *Infect Immun* **64**, 154-160.

Prasadarao N. V., Wass C. A., Stins M. F., Shimada H., Kim K. S. (1999). Outer membrane protein A-promoted actin condensation of brain microvascular endothelial cells is required for *Escherichia coli* invasion. *Infect Immun* **67**, 5775-5783.

Pratt L. A. & Kolter R. (1998). Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* **30**, 285-293.

Prigent-Combaret C., Vidal O., Dorel C., Lejeune P. (1999). Abiotic surface sensing and biofilm-dependent regulation of gene expression in *Escherichia coli*. *J Bacteriol* **181**, 5993-6002.

Prigent-Combaret C., Prensier G., Le Thi T. T., Vidal O., Lejeune P., Dorel C. (2000). Developmental pathway for biofilm formation in curli-producing *Escherichia coli* strains: role of flagella, curli and colanic acid. *Environ Microbiol* **2**, 450-464.

Proft T. & Baker E. N. (2009). Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. *Cell Mol Life Sci* **66**, 613-635.

Prüß B. M., Besemann C., Denton A., Wolfe A. J. (2006). A complex transcription network controls the early stages of biofilm development by *Escherichia coli*. *J Bacteriol* **188**, 3731-3739.

Pupo G. M., Lan R., Reeves P. R. (2000). Multiple independent origins of Shigella clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proc Natl Acad Sci U S A* **97**, 10567-10572.

Pupo G. M., Karaolis D. K., Lan R., Reeves P. R. (1997). Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. *Infect Immun* **65**, 2685-2692.

Puttamreddy S., Cornick N. A., Minion F. C. (2010). Genome-wide transposon mutagenesis reveals a role for pO157 genes in biofilm development in *Escherichia coli* O157:H7 EDL933. *Infect Immun* **78**, 2377-2384.

Py B., Causton H., Mudd E. A., Higgins C. F. (1994). A protein complex mediating mRNA degradation in *Escherichia coli*. *Mol Microbiol* **14**, 717-729.

Qin J., Li R., Raes J., Arumugam M., Burgdorf K. S., Manichanh C., Nielsen T., Pons N., Levenez F.& other authors. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65.

Rasko D. A., Rosovitz M. J., Myers G. S., Mongodin E. F., Fricke W. F., Gajer P., Crabtree J., Sebaihia M., Thomson N. R.& other authors. (2008). The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J Bacteriol* **190**, 6881-6893.

Ray K., Marteyn B., Sansonetti P. J., Tang C. M. (2009). Life on the inside: the intracellular lifestyle of cytosolic bacteria. *Nat Rev Microbiol* **7**, 333-340.

Régnier P. & Grunberg-Manago M. (1989). Cleavage by RNase III in the transcripts of the *metY-nusA-infB* operon of *Escherichia coli* releases the tRNA and initiates the decay of the downstream mRNA. *J Mol Biol* **210**, 293-302.

Reid S. D., Herbelin C. J., Bumbaugh A. C., Selander R. K., Whittam T. S. (2000). Parallel evolution of virulence in pathogenic *Escherichia coli* . *Nature* **406**, 64-67.

Reisner A., Haagensen J. A., Schembri M. A., Zechner E. L., Molin S. (2003). Development and maturation of *Escherichia coli* K-12 biofilms. *Mol Microbiol* **48**, 933-946.

Reiss D. J. & Mobley H. L. (2011). Determination of the target sequence bound by PapX, a repressor of bacterial motility, in the *flhD* promoter using SELEX and high-throughput sequencing. *J Biol Chem* .

Rendón M. A., Saldaña Z., Erdem A. L., Monteiro-Neto V., Vázquez A., Kaper J. B., Puente J. L., Girón J. A. (2007). Commensal and pathogenic *Escherichia coli* use a common pilus adherence factor for epithelial cell colonization. *Proc Natl Acad Sci U S A* **104**, 10637-10642.

Repoila F. & Gottesman S. (2001). Signal transduction cascade for regulation of RpoS: temperature regulation of DsrA. *J Bacteriol* **183**, 4012-4023.

Richardson J. P., Grimley C., Lowery C. (1975). Transcription termination factor Rho activity is altered in *Escherichia coli* with *suA* gene mutations. *Proc Natl Acad Sci U S A* **72**, 1725-1728.

Robbins J. B., McCracken G. H.,Jr, Gotschlich E. C., Ørskov F., Ørskov I., Hanson L. A. (1974). *Escherichia coli* K1 capsular polysaccharide associated with neonatal meningitis. *N Engl J Med* **290**, 1216-1220.

Roberts J. W. (1969). Termination factor for RNA synthesis. *Nature* **224**, 1168-1174.

Ron E. Z., Yerushalmi Z., Naveh M. W. (1991). Adherence pili of *Escherichia coli* O78 determine host and tissue specificity. In *Microbial Surface Components and Toxins in Relation to Pathogenesis,* pp. 61-68. Edited by E. Z. Ron & S. Rottem. New York: Plenum Press.

Rose M. & Charlesworth B. (1980). A test of evolutionary theories of senescence. *Nature* **287**, 141-142.

Rotimi V. O. & Duerden B. I. (1981). The development of the bacterial flora in normal neonates. *J Med Microbiol* **14**, 51-62.

Rowley G., Spector M., Kormanec J., Roberts M. (2006). Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nat Rev Microbiol* **4**, 383-394.

Roy K., Hilliard G. M., Hamilton D. J., Luo J., Ostmann M. M., Fleckenstein J. M. (2009). Enterotoxigenic *Escherichia coli* EtpA mediates adhesion between flagella and host cells. *Nature* **457**, 594-598.

Rubin L. L. & Staddon J. M. (1999). The cell biology of the blood-brain barrier. *Annu Rev Neurosci* **22**, 11-28.

Saarela S., Westerlund-Wikström B., Rhen M., Korhonen T. K. (1996). The GafD protein of the G (F17) fimbrial complex confers adhesiveness of *Escherichia coli* to laminin. *Infect Immun* **64**, 2857-2860.

Saini S., Slauch J. M., Aldridge P. D., Rao C. V. (2010). Role of cross talk in regulating the dynamic expression of the flagellar *Salmonella* pathogenicity island 1 and type 1 fimbrial genes. *J Bacteriol* **192**, 5767-5777.

Sakellaris H., Hannink N. K., Rajakumar K., Bulach D., Hunt M., Sasakawa C., Adler B. (2000). Curli loci of *Shigella* spp. *Infect Immun* **68**, 3780-3783.

Saldaña Z., Erdem A. L., Schüller S., Okeke I. N., Lucas M., Sivananthan A., Phillips A. D., Kaper J. B., Puente J. L., Girón J. A. (2009). The *Escherichia coli* common pilus and the bundle-forming pilus act in concert during the formation of localized adherence by enteropathogenic *E. coli*. *J Bacteriol* **191**, 3451-3461.

Salit I. E. & Gotschlich E. C. (1977). Type I *Escherichia coli* pili: characterization of binding to monkey kidney cells. *J Exp Med* **146**, 1182-1194.

Samuel G. & Reeves P. (2003). Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr Res* **338**, 2503-2519.

Sands J. F., Régnier P., Cummings H. S., Grunberg-Manago M., Hershey J. W. (1988). The existence of two genes between *infB* and *rpsO* in the *Escherichia coli* genome: DNA sequencing and S1 nuclease mapping. *Nucleic Acids Res* **16**, 10803-10816.

Sarén A., Virkola R., Hacker J., Korhonen T. K. (1999). The cellular form of human fibronectin as an adhesion target for the S fimbriae of meningitis-associated *Escherichia coli*. *Infect Immun* **67**, 2671-2676.

Sarff L. D., McCracken G. H., Schiffer M. S., Glode M. P., Robbins J. B., Ørskov I., Ørskov F. (1975). Epidemiology of *Escherichia coli* K1 in healthy and diseased newborns. *Lancet* **1**, 1099-1104.

Sauter S. L., Rutherfurd S. M., Wagener C., Shively J. E., Hefta S. A. (1991). Binding of nonspecific cross-reacting antigen, a granulocyte membrane glycoprotein, to *Escherichia coli* expressing type 1 fimbriae. *Infect Immun* **59**, 2485-2493.

Savage D. C. (1977). Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol* **31**, 107-133.

Savageau M.A. (1983). *Escherichia coli* habitats, cell types, and molecular mechanisms of gene control. *Am Nat* **122**, 732-744.

Scaletsky I. C., Aranda K. R., Souza T. B., Silva N. P. (2010). Adherence factors in atypical enteropathogenic *Escherichia coli* strains expressing the localized adherence-like pattern in HEp-2 cells. *J Clin Microbiol* **48**, 302-306.

Schamberger G. P., Phillips R. L., Jacobs J. L., Diez-Gonzalez F. (2004). Reduction of *Escherichia coli* O157:H7 populations in cattle by addition of colicin E7-producing *E. coli* to feed. *Appl Environ Microbiol* **70**, 6053-6060.

Schembri M. A. & Klemm P. (2001). Biofilm formation in a hydrodynamic environment by novel FimH variants and ramifications for virulence. *Infect Immun* **69**, 1322-1328.

Schembri M. A., Dalsgaard D., Klemm P. (2004). Capsule shields the function of short bacterial adhesins. *J Bacteriol* **186**, 1249-1257.

Schembri M. A., Kjærgaard K., Klemm P. (2003). Global gene expression in *Escherichia coli* biofilms. *Mol Microbiol* **48**, 253-267.

Schmöe K., Rogov V. V., Rogova N. Y., Löhr F., Güntert P., Bernhard F., Dötsch V. (2011). Structural insights into Rcs phosphotransfer: the newly identified RcsD-ABL domain enhances interaction with the response regulator RcsB. *Structure* **19**, 577-587.

Schmoll T., Ott M., Oudega B., Hacker J. (1990). Use of a wild-type gene fusion to determine the influence of environmental conditions on expression of the S fimbrial adhesin in an *Escherichia coli* pathogen. *J Bacteriol* **172**, 5103-5111.

Schultz M. (2008). Clinical use of *E. coli* Nissle 1917 in inflammatory bowel disease. *Inflamm Bowel Dis* **14**, 1012-1018.

Schwan W. R., Shibata S., Aizawa S., Wolfe A. J. (2007). The two-component response regulator RcsB regulates type 1 piliation in *Escherichia coli*. *J Bacteriol* **189**, 7159-7163.

Schwan W. R., Lee J. L., Lenard F. A., Matthews B. T., Beck M. T. (2002). Osmolarity and pH growth conditions regulate *fim* gene transcription and type 1 pilus expression in uropathogenic *Escherichia coli*. *Infect Immun* **70**, 1391-1402.

Scott J. R. & Zähner D. (2006). Pili with strong attachments: Gram-positive bacteria do it differently. *Mol Microbiol* **62**, 320-330.

Selander R. K. & Levin B. R. (1980). Genetic diversity and structure in *Escherichia coli* populations. *Science* **210**, 545-547.

Selander R. K., Korhonen T. K., Väisanen-Rhen V., Williams P. H., Pattison P. E., Caugant D. A. (1986). Genetic relationships and clonal structure of strains of *Escherichia coli* causing neonatal septicemia and meningitis. *Infect Immun* **52**, 213-222.

Selvarangan R., Goluszko P., Singhal J., Carnoy C., Moseley S., Hudson B., Nowicki S., Nowicki B. (2004). Interaction of Dr adhesin with collagen type IV is a critical step in *Escherichia coli* renal persistence. *Infect Immun* **72**, 4827-4835.

Serganov A. & Patel D. J. (2007). Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet* **8**, 776-790.

Sheikh J., Hicks S., Dall'Agnol M., Phillips A. D., Nataro J. P. (2001). Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Mol Microbiol* **41**, 983-997.

Shimada T., Fujita N., Yamamoto K., Ishihama A. (2011). Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS One* **6**, e20081.

Shin S. & Park C. (1995). Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR. *J Bacteriol* **177**, 4696-4702.

Shipley P. L., Gyles C. L., Falkow S. (1978). Characterization of plasmids that encode for the K88 colonization antigen. *Infect Immun* **20**, 559-566.

Silverman M. & Simon M. (1974). Characterization of *Escherichia coli* flagellar mutants that are insensitive to catabolite repression. *J Bacteriol* **120**, 1196-1203.

Simm R., Morr M., Kader A., Nimtz M., Römling U. (2004). GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* **53**, 1123-1134.

Simms A. N. & Mobley H. L. (2008). PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic *Escherichia coli*. *Infect Immun* **76**, 4833-4841.

Simon R., Priefer U., Pühler A. (1983). A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria *Biotechnology* **1**, 784-791.

Simons R. W., Houman F., Kleckner N. (1987). Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**, 85-96.

Sims G. E. & Kim S. H. (2011). Whole-genome phylogeny of *Escherichia coli*/*Shigella* group by feature frequency profiles (FFPs). *Proc Natl Acad Sci U S A* **108**, 8329-8334.

Sjöbring U., Pohl G., Olsén A. (1994). Plasminogen, absorbed by *Escherichia coli* expressing curli or by *Salmonella enteritidis* expressing thin aggregative fimbriae, can be activated by simultaneously captured tissue-type plasminogen activator (t-PA). *Mol Microbiol* **14**, 443-452.

Sjöström A. E., Balsalobre C., Emödy L., Westerlund-Wikström B., Hacker J., Uhlin B. E. (2009a). The SfaXII protein from newborn meningitis *E. coli* is involved in regulation of motility and type 1 fimbriae expression. *Microb Pathog* **46**, 243-252.

Sjöström A. E., Sondén B., Müller C., Rydström A., Dobrindt U., Wai S. N., Uhlin B. E. (2009b). Analysis of the *sfaX*$_{II}$ locus in the *Escherichia coli* meningitis isolate IHE3034 reveals two novel regulatory genes within the promoter-distal region of the main S fimbrial operon. *Microb Pathog* **46**, 150-158.

Sledjeski D. D., Gupta A., Gottesman S. (1996). The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J* **15**, 3993-4000.

Smits W. K., Kuipers O. P., Veening J. W. (2006). Phenotypic variation in bacteria: the role of feedback regulation. *Nat Rev Microbiol* **4**, 259-271.

Snyder J. A., Haugen B. J., Lockatell C. V., Maroncle N., Hagan E. C., Johnson D. E., Welch R. A., Mobley H. L. (2005). Coordinate expression of fimbriae in uropathogenic *Escherichia coli*. *Infect Immun* **73**, 7588-7596.

Sokurenko E. V., Courtney H. S., Maslow J., Siitonen A., Hasty D. L. (1995). Quantitative differences in adhesiveness of type 1 fimbriated *Escherichia coli* due to structural differences in *fimH* genes. *J Bacteriol* **177**, 3680-3686.

Sokurenko E. V., Courtney H. S., Ohman D. E., Klemm P., Hasty D. L. (1994). FimH family of type 1 fimbrial adhesins: functional heterogeneity due to minor sequence variations among *fimH* genes. *J Bacteriol* **176**, 748-755.

Sokurenko E. V., Chesnokova V., Dykhuizen D. E., Ofek I., Wu X. R., Krogfelt K. A., Struve C., Schembri M. A., Hasty D. L. (1998). Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. *Proc Natl Acad Sci U S A* **95**, 8922-8926.

Sorek R. & Cossart P. (2010). Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet* **11**, 9-16.

Soto G. E. & Hultgren S. J. (1999). Bacterial adhesins: common themes and variations in architecture and assembly. *J Bacteriol* **181**, 1059-1071.

Soutourina O., Kolb A., Krin E., Laurent-Winter C., Rimsky S., Danchin A., Bertin P. (1999). Multiple control of flagellum biosynthesis in *Escherichia coli*: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the *flhDC* master operon. *J Bacteriol* **181**, 7500-7508.

Soutourina O. A. & Bertin P. N. (2003). Regulation cascade of flagellar expression in Gram-negative bacteria. *FEMS Microbiol Rev* **27**, 505-523.

Stark P. L. & Lee A. (1982). The microbial ecology of the large bowel of breast-fed and formula-fed infants during the first year of life. *J Med Microbiol* **15**, 189-203.

Stevenson G., Neal B., Liu D., Hobbs M., Packer N. H., Batley M., Redmond J. W., Lindquist L., Reeves P. (1994). Structure of the O antigen of *Escherichia coli* K-12 and the sequence of its *rfb* gene cluster. *J Bacteriol* **176**, 4144-4156.

Stewart P. S. & Franklin M. J. (2008). Physiological heterogeneity in biofilms. *Nat Rev Microbiol* **6**, 199-210.

Stins M. F., Nemani P. V., Wass C., Kim K. S. (1999). *Escherichia coli* binding to and invasion of brain microvascular endothelial cells derived from humans and rats of different ages. *Infect Immun* **67**, 5522-5525.

Stoebel D. M., Free A., Dorman C. J. (2008). Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154**, 2533-2545.

Stone K. D., Zhang H. Z., Carlson L. K., Donnenberg M. S. (1996). A cluster of fourteen genes from enteropathogenic *Escherichia coli* is sufficient for the biogenesis of a type IV pilus. *Mol Microbiol* **20**, 325-337.

Stoodley P., Sauer K., Davies D. G., Costerton J. W. (2002). Biofilms as complex differentiated communities. *Annu Rev Microbiol* **56**, 187-209.

Storz G., Vogel J., Wassarman K. M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* **43**, 880-891.

Strömberg N., Marklund B. I., Lund B., Ilver D., Hamers A., Gaastra W., Karlsson K. A., Normark S. (1990). Host-specificity of uropathogenic *Escherichia coli* depends on differences in binding specificity to Galα1-4Gal-containing isoreceptors. *EMBO J* **9**, 2001-2010.

Studier F. W. & Moffatt B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-130.

Sukumaran S. K., Selvaraj S. K., Prasadarao N. V. (2004). Inhibition of apoptosis by *Escherichia coli* K1 is accompanied by increased expression of Bcl_{XL} and blockade of mitochondrial cytochrome *c* release in macrophages. *Infect Immun* **72**, 6012-6022.

Sullivan T. D., LaScolea L. J.,Jr, Neter E. (1982). Relationship between the magnitude of bacteremia in children and the clinical disease. *Pediatrics* **69**, 699-702.

Svensson A., Larsson A., Emtenäs H., Hedenström M., Fex T., Hultgren S. J., Pinkner J. S., Almqvist F., Kihlberg J. (2001). Design and evaluation of pilicides: potential novel antibacterial agents directed against uropathogenic *Escherichia coli*. *Chembiochem* **2**, 915-918.

Swidsinski A., Mendling W., Loening-Baucke V., Ladhoff A., Swidsinski S., Hale L. P., Lochs H. (2005). Adherent biofilms in bacterial vaginosis. *Obstet Gynecol* **106**, 1013-1023.

Takeda S., Fujisawa Y., Matsubara M., Aiba H., Mizuno T. (2001). A novel feature of the multistep phosphorelay in *Escherichia coli*: a revised model of the RcsC → YojN → RcsB signalling pathway implicated in capsular synthesis and swarming behaviour. *Mol Microbiol* **40**, 440-450.

Tenaillon O., Skurnik D., Picard B., Denamur E. (2010). The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* **8**, 207-217.

Teng C. H., Cai M., Shin S., Xie Y., Kim K. J., Khan N. A., Di Cello F., Kim K. S. (2005). *Escherichia coli* K1 RS218 interacts with human brain microvascular endothelial cells via type 1 fimbria bacteria in the fimbriated state. *Infect Immun* **73**, 2923-2931.

Thanassi D. G., Cheng L. W., Nikaido H. (1997). Active efflux of bile salts by *Escherichia coli*. *J Bacteriol* **179**, 2512-2518.

Thomas W. E., Nilsson L. M., Forero M., Sokurenko E. V., Vogel V. (2004). Shear-dependent 'stick-and-roll' adhesion of type 1 fimbriated *Escherichia coli*. *Mol Microbiol* **53**, 1545-1557.

Thomas W. E., Trintchina E., Forero M., Vogel V., Sokurenko E. V. (2002). Bacterial adhesion to target cells enhanced by shear force. *Cell* **109**, 913-923.

Tivendale K. A., Logue C. M., Kariyawasam S., Jordan D., Hussein A., Li G., Wannemuehler Y., Nolan L. K. (2010). Avian-pathogenic *Escherichia coli strains* are similar to neonatal meningitis *E. coli* strains and are able to cause meningitis in the rat model of human disease. *Infect Immun* **78**, 3412-3419.

Tobe T., Nakanishi N., Sugimoto N. (2011). Activation of motility by sensing short-chain fatty acids via two steps in a flagellar gene regulatory cascade in enterohemorrhagic *Escherichia coli*. *Infect Immun* **79**, 1016-1024.

Tomoyasu T., Takaya A., Isogai E., Yamamoto T. (2003). Turnover of FlhD and FlhC, master regulator proteins for *Salmonella* flagellum biogenesis, by the ATP-dependent ClpXP protease. *Mol Microbiol* **48**, 443-452.

Totsika M., Beatson S. A., Holden N., Gally D. L. (2008). Regulatory interplay between *pap* operons in uropathogenic *Escherichia coli*. *Mol Microbiol* **67**, 996-1011.

Touchon M., Hoede C., Tenaillon O., Barbe V., Baeriswyl S., Bidet P., Bingen E., Bonacorsi S., Bouchier C.& other authors. (2009). Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**, e1000344.

Troy F. A.,2nd. (1992). Polysialylation: from bacteria to brains. *Glycobiology* **2**, 5-23.

Turnbaugh P. J., Hamady M., Yatsunenko T., Cantarel B. L., Duncan A., Ley R. E., Sogin M. L., Jones W. J., Roe B. A.& other authors. (2009). A core gut microbiome in obese and lean twins. *Nature* **457**, 480-484.

Turner S. M., Chaudhuri R. R., Jiang Z. D., DuPont H., Gyles C., Penn C. W., Pallen M. J., Henderson I. R. (2006). Phylogenetic comparisons reveal multiple acquisitions of the toxin genes by enterotoxigenic *Escherichia coli* strains of different evolutionary lineages. *J Clin Microbiol* **44**, 4528-4536.

Uhlich G. A., Gunther N. W.,4th, Bayles D. O., Mosier D. A. (2009). The CsgA and Lpp proteins of an *Escherichia coli* O157:H7 strain affect HEp-2 cell invasion, motility, and biofilm formation. *Infect Immun* **77**, 1543-1552.

Unhanand M., Mustafa M. M., McCracken G. H.,Jr, Nelson J. D. (1993). Gram-negative enteric bacillary meningitis: a twenty-one-year experience. *J Pediatr* **122**, 15-21.

Valle J., Da Re S., Henry N., Fontaine T., Balestrino D., Latour-Lambert P., Ghigo J. M. (2006). Broad-spectrum biofilm inhibition by a secreted bacterial polysaccharide. *Proc Natl Acad Sci U S A* **103**, 12558-12563.

van der Woude M. W. & Low D. A. (1994). Leucine-responsive regulatory protein and deoxyadenosine methylase control the phase variation and expression of the *sfa* and *daa* pili operons in *Escherichia coli*. *Mol Microbiol* **11**, 605-618.

Van Loy C. P., Sokurenko E. V., Moseley S. L. (2002). The major structural subunits of Dr and F1845 fimbriae are adhesins. *Infect Immun* **70**, 1694-1702.

Vejborg R. M., Hancock V., Petersen A. M., Krogfelt K. A., Klemm P. (2011). Comparative genomics of *Escherichia coli* isolated from patients with inflammatory bowel disease. *BMC Genomics* **12**, 316.

Vejborg R. M., Friis C., Hancock V., Schembri M. A., Klemm P. (2010). A virulent parent with probiotic progeny: comparative genomics of *Escherichia coli* strains CFT073, Nissle 1917 and ABU 83972. *Mol Genet Genomics* **283**, 469-484.

Venkatesan M. M., Goldberg M. B., Rose D. J., Grotbeck E. J., Burland V., Blattner F. R. (2001). Complete DNA sequence and analysis of the large virulence plasmid of *Shigella flexneri*. *Infect Immun* **69**, 3271-3285.

Vianney A., Jubelin G., Renault S., Dorel C., Lejeune P., Lazzaroni J. C. (2005). *Escherichia coli tol* and *rcs* genes participate in the complex network affecting curli synthesis. *Microbiology* **151**, 2487-2497.

Vidal O., Longin R., Prigent-Combaret C., Dorel C., Hooreman M., Lejeune P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new *ompR* allele that increases curli expression. *J Bacteriol* **180**, 2442-2449.

Virkola R. (1987). Binding characteristics of *Escherichia coli* type 1 fimbriae in the human kidney. *FEMS Microbiol Lett* **40**, 257-262.

Virkola R., Parkkinen J., Hacker J., Korhonen T. K. (1993). Sialyloligosaccharide chains of laminin as an extracellular matrix target for S fimbriae of *Escherichia coli*. *Infect Immun* **61**, 4480-4484.

Waksman G. & Hultgren S. J. (2009). Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat Rev Microbiol* **7**, 765-774.

Walk S. T., Alm E. W., Calhoun L. M., Mladonicky J. M., Whittam T. S. (2007). Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environ Microbiol* **9**, 2274-2288.

Walk S. T., Alm E. W., Gordon D. M., Ram J. L., Toranzos G. A., Tiedje J. M., Whittam T. S. (2009). Cryptic lineages of the genus *Escherichia* . *Appl Environ Microbiol* **75**, 6534-6544.

Wang X., Rochon M., Lamprokostopoulou A., Lünsdorf H., Nimtz M., Römling U. (2006). Impact of biofilm matrix components on interaction of commensal *Escherichia coli* with the gastrointestinal cell line HT-29. *Cell Mol Life Sci* **63**, 2352-2363.

Wang Y. & Kim K. S. (2000). Effect of *rpoS* mutations on stress-resistance and invasion of brain microvascular endothelial cells in *Escherichia coli* K1. *FEMS Microbiol Lett* **182**, 241-247.

Warren J. W. (1996). Clinical presentation and epidemiology of urinary tract infections. In *Urinary Tract Infections: Molecular Pathogenesis and Clinical Management,* pp. 3-27. Edited by H. L. T. Mobley & J. W. Warren. Washington, DC.: ASM Press.

Waters L. S. & Storz G. (2009). Regulatory RNAs in bacteria. *Cell* **136**, 615-628.

Watt S., Lanotte P., Mereghetti L., Moulin-Schouleur M., Picard B., Quentin R. (2003). *Escherichia coli* strains from pregnant women and neonates: intraspecies genetic distribution and prevalence of virulence factors. *J Clin Microbiol* **41**, 1929-1935.

Wehland M. & Bernhard F. (2000). The RcsAB box. Characterization of a new operator essential for the regulation of exopolysaccharide biosynthesis in enteric bacteria. *J Biol Chem* **275**, 7013-7020.

Wei B. L., Brun-Zinkernagel A. M., Simecka J. W., Prüß B. M., Babitzke P., Romeo T. (2001). Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*. *Mol Microbiol* **40**, 245-256.

Weissman S. J., Beskhlebnaya V., Chesnokova V., Chattopadhyay S., Stamm W. E., Hooton T. M., Sokurenko E. V. (2007). Differential stability and trade-off effects of pathoadaptive mutations in the *Escherichia coli* FimH adhesin. *Infect Immun* **75**, 3548-3555.

Weissman S. J., Chattopadhyay S., Aprikian P., Obata-Yasuoka M., Yarova-Yarovaya Y., Stapleton A., Ba-Thein W., Dykhuizen D., Johnson J. R., Sokurenko E. V. (2006). Clonal analysis reveals high rate of structural mutations in fimbrial adhesins of extraintestinal pathogenic *Escherichia coli*. *Mol Microbiol* **59**, 975-988.

Welch R. A., Burland V., Plunkett G.,3rd, Redford P., Roesch P., Rasko D., Buckles E. L., Liou S. R., Boutin A.& other authors. (2002). Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A* **99**, 17020-17024.

Westerlund B. & Korhonen T. K. (1993). Bacterial proteins binding to the mammalian extracellular matrix. *Mol Microbiol* **9**, 687-694.

Westerlund B., Kuusela P., Vartio T., van Die I., Korhonen T. K. (1989a). A novel lectin-independent interaction of P fimbriae of *Escherichia coli* with immobilized fibronectin. *FEBS Lett* **243**, 199-204.

Westerlund B., Kuusela P., Risteli J., Risteli L., Vartio T., Rauvala H., Virkola R., Korhonen T. K. (1989b). The O75X adhesin of uropathogenic *Escherichia coli* is a type IV collagen-binding protein. *Mol Microbiol* **3**, 329-337.

Westerlund B., van Die I., Kramer C., Kuusela P., Holthöfer H., Tarkkanen A. -., Virkola R., Riegman N., Bergmans H., Hoekstra W. (1991). Multifunctional nature of P fimbriae of uropathogenic *Escherichia coli*: mutations in *fsoE* and *fsoF* influence fimbrial binding to renal tubuli and immobilized fibronectin. *Mol Microbiol* **5**, 2965-2975.

Weyand N. J., Braaten B. A., van der Woude M., Tucker J., Low D. A. (2001). The essential role of the promoter-proximal subunit of CAP in *pap* phase variation: Lrp- and helical phase-dependent activation of *papBA* transcription by CAP from -215. *Mol Microbiol* **39**, 1504-1522.

White A. P., Sibley K. A., Sibley C. D., Wasmuth J. D., Schaefer R., Surette M. G., Edge T. A., Neumann N. F. (2011). Intergenic sequence comparison of *Escherichia coli* isolates reveals lifestyle adaptations but not host specificity. *Appl Environ Microbiol* **77**, 7620-7632.

White-Ziegler C. A., Malhowski A. J., Young S. (2007). Human body temperature (37°C) increases the expression of iron, carbohydrate, and amino acid utilization genes in *Escherichia coli* K-12. *J Bacteriol* **189**, 5429-5440.

White-Ziegler C. A., Um S., Pérez N. M., Berns A. L., Malhowski A. J., Young S. (2008). Low temperature (23 °C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in *Escherichia coli* K-12. *Microbiology* **154**, 148-166.

Wilson K. S. & von Hippel P. H. (1995). Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci U S A* **92**, 8793-8797.

Wirth T., Falush D., Lan R., Colles F., Mensa P., Wieler L. H., Karch H., Reeves P. R., Maiden M. C., Ochman H., Achtman M. (2006). Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* **60**, 1136-1151.

Wood T. K., Gonzalez Barrios A. F., Herzberg M., Lee J. (2006). Motility influences biofilm architecture in *Escherichia coli*. *Appl Microbiol Biotechnol* **72**, 361-367.

Wooster D. G., Maruvada R., Blom A. M., Prasadarao N. V. (2006). Logarithmic phase *Escherichia coli* K1 efficiently avoids serum killing by promoting C4bp-mediated C3b and C4b degradation. *Immunology* **117**, 482-493.

Wright K. J., Seed P. C., Hultgren S. J. (2007). Development of intracellular bacterial communities of uropathogenic *Escherichia coli* depends on type 1 pili. *Cell Microbiol* **9**, 2230-2241.

Wright K. J., Seed P. C., Hultgren S. J. (2005). Uropathogenic *Escherichia coli* flagella aid in efficient urinary tract colonization. *Infect Immun* **73**, 7657-7668.

Xia Y., Gally D., Forsman-Semb K., Uhlin B. E. (2000). Regulatory cross-talk between adhesin operons in *Escherichia coli*: inhibition of type 1 fimbriae expression by the PapB protein. *EMBO J* **19**, 1450-1457.

Xicohtencatl-Cortes J., Monteiro-Neto V., Saldaña Z., Ledesma M. A., Puente J. L., Girón J. A. (2009). The type 4 pili of enterohemorrhagic *Escherichia coli* O157:H7 are multipurpose structures with pathogenic attributes. *J Bacteriol* **191**, 411-421.

Xue P., Corbett D., Goldrick M., Naylor C., Roberts I. S. (2009). Regulation of expression of the region 3 promoter of the *Escherichia coli* K5 capsule gene cluster involves H-NS, SlyA, and a large 5' untranslated region. *J Bacteriol* **191**, 1838-1846.

Yamamoto K., Hirao K., Oshima T., Aiba H., Utsumi R., Ishihama A. (2005). Functional characterization *in vitro* of all two-component signal transduction systems from *Escherichia coli*. *J Biol Chem* **280**, 1448-1456.

Yang J., Tauschek M., Strugnell R., Robins-Browne R. M. (2005). The H-NS protein represses transcription of the *eltAB* operon, which encodes heat-labile enterotoxin in enterotoxigenic *Escherichia coli*, by binding to regions downstream of the promoter. *Microbiology* **151**, 1199-1208.

Yanisch-Perron C., Vieira J., Messing J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

Yao Y., Xie Y., Perace D., Zhong Y., Lu J., Tao J., Guo X., Kim K. S. (2009). The type III secretion system is involved in the invasion and intracellular survival of *Escherichia coli* K1 in human brain microvascular endothelial cells. *FEMS Microbiol Lett* **300**, 18-24.

Zdziarski J., Svanborg C., Wullt B., Hacker J., Dobrindt U. (2008). Molecular basis of commensalism in the urinary tract: low virulence or virulence attenuation? *Infect Immun* **76**, 695-703.

Zelmer A., Bowen M., Jokilammi A., Finne J., Luzio J. P., Taylor P. W. (2008). Differential expression of the polysialyl capsule during blood-to-brain transit of neuropathogenic *Escherichia coli* K1. *Microbiology* **154**, 2522-2532.

Zhang L., Foxman B., Marrs C. (2002). Both urinary and rectal *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. *J Clin Microbiol* **40**, 3951-3955.

Zheng D., Constantinidou C., Hobman J. L., Minchin S. D. (2004). Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res* **32**, 5874-5893.

Zhou G., Mo W. J., Sebbel P., Min G., Neubert T. A., Glockshuber R., Wu X. R., Sun T. T., Kong X. P. (2001). Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from in vitro FimH binding. *J Cell Sci* **114**, 4095-4103.